OM nucleic - nucleic search, using sw model

Run on: May 5, 2004, 09:55:26 ; Search time 4891.91 Seconds
                                      (without alignments)
                                      11571.357 Million cell updates/sec

Title: US-09-901-572A-1
Perfect score: 1306
Sequence: 1 aaaaacatcagattgttaat........taaaatcgtttatcaggtt 1306

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699 5 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
               1: gb_ba:*
               2: gb_htg:*
               3: gb_in:*
               4: gb_om:*
               5: gb_ov:*
               6: gb_pat:*
               7: gb_ph:*
               8: gb_pl:*
               9: gb_pr:*
              10: gb_ro:*
              11: gb_sts:*
              12: gb_sy:*
              13: gb_un:*
              14: gb_vi:*
              15: em_ba:*
              16: em_fun:*
              17: em_hum:*
              18: em_in:*
              19: em_mu:*
              20: em_om:*
              21: em_or:*
              22: em_ov:*
              23: em_pat:*
              24: em_ph:*
              25: em_pl:*
              26: em_ro:*
              27: em_sts:*
              28: em_un:*
              29: em_vi:*
              30: em_htg_hum:*
              31: em_htg_inv:*
              32: em_htg_other:*
              33: em_htg_mus:*
              34: em_htg_pln:*
              35: em_htg_rod:*
              36: em_htg_mam:*
              37: em_htg_vrt:*
              38: em_sy:*
              39: em_htgo_hum:*
              40: em_htgo_mus:*
              41: em_htgo_other:*

          GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1280.4	98.0	2144	6	AR035279	AR035279 Sequence
2	1270	97.2	1387	6	AR035276	AR035276 Sequence
3	1264	96.8	1305	6	I17387	I17387 Sequence 1
4	1262.4	96.7	1387	6	E09301	E09301 DNA encodin
5	1179.4	90.3	302070	1	AB016968	AB016968 Mycoplasm
6	1086.8	83.2	2014	6	AR035278	AR035278 Sequence
7	1028	78.7	1152	6	AX665164	AX665164 Sequence
8	989.4	75.8	1015	6	E09896	E09896 Mycoplasma
9	986.2	75.5	1015	6	S65869	S65869 TM-1=29 kda
10	921.6	70.6	1082	6	AX665187	AX665187 Sequence
11	824.2	63.1	853	6	E02348	E02348 DNA sequenc
12	679.8	52.1	708	6	E02342	E02342 DNA sequenc
13	250.6	19.2	301042	1	AB016967	AB016967 Mycoplasm
14	249	19.1	4568	1	AF210770	AF210770 Mycoplasm
c 15	222.8	17.1	301901	1	AB016969	AB016969 Mycoplasm
16	217	16.6	7434	1	AF275312	AF275312 Mycoplasm
17	216.6	16.6	2369	6	AR035275	AR035275 Sequence
18	216.6	16.6	2369	6	E09895	E09895 Mycoplasma
19	215	16.5	3331	1	AF141940	AF141940 Mycoplasm
20	206.2	15.8	9720	1	MYCPMGAB	L28424 Mycoplasma
c 21	203.8	15.6	302070	1	AB016968	AB016968 Mycoplasm
22	199.8	15.3	3894	1	AY065985	AY065985 Mycoplasm
23	193.4	14.8	4715	1	AF032890	AF032890 Mycoplasm
24	191.8	14.7	8904	1	MGU90714	U90714 Mycoplasma
25	190.2	14.6	3897	1	AF053978	AF053978 Mycoplasm
26	173.2	13.3	2803	1	MYCHAHOM	M83178 Mycoplasma
27	173.2	13.3	2809	1	S55216	S55216 pMGA1.2=hem
28	136.4	10.4	702	6	E02344	E02344 DNA sequence
c 29	92	7.0	170627	2	AC125567	AC125567 Rattus no
30	89	6.8	146275	2	AL935272	AL935272 Danio rer
31	88.2	6.8	116807	2	BX890614	BX890614 Danio rer
c 32	85	6.5	110000	3	AC116305_1	Continuation (2 of
c 33	84.6	6.5	198743	2	BX530070	BX530070 Danio rer
34	82.4	6.3	94534	2	AL929250	AL929250 Zebrafish
c 35	82.2	6.3	155204	2	AC007926	AC007926 Trypanoso
36	80	6.1	154071	3	AC115598	AC115598 Dictyoste
37	79.2	6.1	153751	3	AC116551	AC116551 Dictyoste
c 38	79.2	6.1	225581	2	BX537105	BX537105 Danio rer
39	78.8	6.0	199347	2	BX321884	BX321884 Danio rer
40	78.6	6.0	213544	2	BX510939	BX510939 Danio rer
41	78.2	6.0	594	6	E02345	E02345 DNA sequenc
42	78.2	6.0	183357	2	BX569779	BX569779 Danio rer
43	77.6	5.9	10593	3	AY232271	AY232271 Dictyoste
44	77.4	5.9	335050	3	PFA929356	AL929356 Plasmodiu
c 45	77.2	5.9	67970	3	PFMAL1P3	AL031746 Plasmodiu

ALIGNMENTS

RESULT 1
AR035279
LOCUS   AR035279   2144 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 9 from patent US 5871742.
ACCESSION   AR035279
VERSION   AR035279.1 GI:5951947
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 2144)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
             Aoyama,S. and Takahashi,K.
  TITLE   Recombinant Avipox virus encoding polypeptide of mycoplasma
             gallisepticum, and utilized a live vaccine

Pred. No. is the number of results predicted by chance to have a

JOURNAL   Patent: US 5871742-A 9 16-FEB-1999;
FEATURES
    source   1..2144
                  /organism="unknown"
                  /mol_type="unassigned DNA"
ORIGIN

Query Match   98.0%;   Score 1280.4;   DB 6;   Length 2144;
Best Local Similarity   98.8%;   Pred. No. 9.2e-177;
Matches 1290;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

```
Qy      1 AAAAACATCAGATTGTAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGTAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAAATAACTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAAATAACTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTAGTCATCTTTTAAGAT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTAGTCATCTTTTAAGAT 180

Qy    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGCTATTACTAAAAAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGCTATTACTAAAAAA 300

Qy    301 GATGCGAAACCCAAATAAATGGCCAAACCCAATTGAGAGCAGCGGCGAATGGAGTTAACAGAT 360
          ||||| |||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db    301 GACGCAAACCCAAATAAATGGCCAAACCCAATTACAAGCAGCGGCGAATGGAGTTAACTGAT 360

Qy    361 CTAATCAATGCTAAAAGCCAAGCCGATGACATTAGCTTCACTACTACAAGAGACATGCCAAGATTGAAGCT 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAAGCCAAGCCGATGACATTAGCTTCACTACTACAAGAGACATGCCAAGATTGAAGCT 420

Qy    421 AGTTTATCATCTGCTTATAGTGAAGTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAAGTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCTAATGAATCAGCCATCAACCAAGCTAATACGGAT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCTAATGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy    541 AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGACACTAAAAACC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGACACTAAAAACC 600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGTTGTCATCAACTGCTTATAATCAAATT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGTTAGCTTCAACTGCTTATAATCAGATT 660

Qy    661 CGCAATAATTAGTGGGATCTATACAAATAAAGCTAGTAGTTAATACTAAAACACTAGAT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTAGTGGGATCTATACAAATAAAGCTAGTAGTTAATACTAAAACACTAGAT 720

Qy    721 CCACTAAATGGGGGGAACGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCACTAAATGGGGGGAATGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy    781 AATAATACGTTATCAACTATTATGACAAAGACAAAAGACTTAATGCATTTACAATCTAATAGT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTATGACAAAGACAAAAGACTTAATGCATTTACAATCTAATAGT 840

Qy    841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy    901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGCTGATGTGATGTAACACCCGTCAAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGCTGATGTGATGTAACACCCGTCAAT 960

Db    901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGCTGATGTGATGTAACACCCGTCAAT 960

Qy    961 TATAAATATGCAAGAAGACCGGTTTGGAATGGTGATGATGGTGAACCTTCAAGTAGAATTCTTGCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TATAAATATGCAAGAAGACCGGTTTGGAATGGTGATGATGGTGAACCTTCAAGTAGAATTCTTGCA 1020

Qy   1021 AACACGAATAGTATCACAGATGTGTTCTTGGATTTATTAGTTAGTGGAACAAACACGAAG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTGTTCTTGGATTTATTAGTTAGTGGGACGAACACGAAG 1080

Qy   1081 TACCAATTTAGTTTTAGCAACATATGGTCCATCAACTCGGTTATTTATATTCCCTTATAAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACCAATTTAGTTTTAGCAACTAACTAACTGGTTATTTATATTCCCTTATAAG 1140

Qy   1141 TTGGTTAAAGCAGTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAAATACTCACAGCTAATCCAACT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAAATACTCACAGCTAATCCAACT 1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
```

RESULT 2
AR035276
LOCUS   AR035276   1387 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION   Sequence 3 from patent US 5871742.
ACCESSION   AR035276
VERSION   AR035276.1   GI:5951944
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 1387)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.
  TITLE   Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 3 16-FEB-1999;
FEATURES   Location/Qualifiers
    source   1..1387
                  /organism="unknown"
                  /mol_type="unassigned DNA"
ORIGIN

Query Match   97.2%;   Score 1270;   DB 6;   Length 1387;
Best Local Similarity   98.8%;   Pred. No. 3.4e-175;
Matches 1290;   Conservative 0;   Mismatches 15;   Indels 1;   Gaps 1;

```
Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA 60
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGT-TAATCTGATATCTTGCTAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAATAACTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAATAACTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTTAAGAT 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTTAAGAT 180

Qy    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTAGT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTAGT 240

Qy    241 TTAGGTACACAACATCCATCCTTTCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTATTACTAAAAAA 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGGTACACAACATCCATCCTTTCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTATTACTAAAAAA 300
```

```
ACCESSION   I17387
VERSION     I17387.1    GI:1597742
KEYWORDS    Unknown.
SOURCE      Unknown.
  ORGANISM  Unclassified.
REFERENCE   1 (bases 1 to 1305)
  AUTHORS   Saito,S., Ohkawa,S., Fujisawa,A., Iritani,Y. and Aoyama,S.
  TITLE     Poultry mycoplasma antigen, gene thereof and recombinant vectors
            containing the gene as well as vaccines utilizing the same
  JOURNAL   Patent: US 5489430-A 1 06-FEB-1996;
FEATURES             Location/Qualifiers
  source    1..1305
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN

Query Match    96.8%;  Score 1264;  DB 6;  Length 1305;
Best Local Similarity  98.3%;  Pred. No. 2.6e-174;
Matches 1284;  Conservative 0;  Mismatches 21;  Indels 1;  Gaps 1;
```

```
Qy     1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAAACACAAAATCTTCTAACAA   60
         |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db     1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAAACACAAAATCTTCTAACAA   60

Qy    61 AATCCTAAATAAAATAAGCCGTTAAAATAACTAACTAAAAAAATAAAAAAAATGGTTTTTCTTATC  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AATCCTAAATAAAATAAGCCGTTAAAATAACTAACTAAAAAAATAAAAAAAATGGTTTTTCTTATC  120

Qy   121 AACCAAAATTCTCTAGTAATAAACGCTATTTATTTTTATTTTTAGTCATCTTTTAAGAT      180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCAAAATTCTCTAGTAATAAACGCTATTTATTTTTATTTTTAGTCATCTTTTAAGAT      180

Qy   181 ATAAATATATCTTAATATCTATGAATGAATAAGAATCATCTTAAAGACTATTAGTTTG      240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTTAATATCTATGAATGAATAAGAATCATCTTAAAGACTATTAGTTTG      240

Qy   241 TTAGTACAACATCCTTCTTAGCATTGGGATTCTAGCTGTGTGTATGTCTATTACTAAAAAAA     300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGTACAACATCCTTCTTAGCATTGGGATTCTAGCTGTGTGTATGTCTATTACTAAAAAAA     300

Qy   301 GATGCAAACCCAATAATGGCCAAACCCAATTACAAGCAGCGCGAATCGCGGCAATGGAGTTAACAGAT 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACGCAAACCCAATAATGGCCAAACCCAATTACAAGCAGCGCGAATCGCGGCAATGGAGTTAACTGAT 360

Qy   361 CTAATCAATGCTAAAGCTGTAAAGGCGATGACATTAGCTTCACTACAAGATTGAAGCT      420
Db   361 CTAATCAATGCTAAAGCTGTAAAGGCGATGACATTAGCTTCACTACAAGATTGAAGCT      420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA     480
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA     480

Qy   481 GAACAACTAAAAATGGCTAAAAATTAGAATTCAGCCATCAACCAAGCTAATACGGAT       540
Db   481 GAACAACTAAAAATGGCTAAAAATTAGAATTCAGCCATCAACCAAGCTAATACGGAT       540

Qy   541 AAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC     600
Db   541 AAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC     600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT    660
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT    660

Qy   661 CGCAATAATTTAGTTGGATCTATACAATAAAGCTAGTTGAATTAATAACTAAAACACTAGAT  720
Db   661 CGTAATAATTTAGTTGGATCTATACAATAAAGCTAGTTGAATTAATAACTAAAACACTAGAT  720

Qy   721 CCACTAAATGGGGAACGCGTTTAGATTCTAATGAGATTTTAGATTCATCTTTAAGAT      780
Db   721 CCACTAAATGGGGAATGCGTTTAGATTCTAATGAGATTTTAGATTCATCTTTAAGAT      780

Qy   781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT   840
Db   781 AATAATACGTTATCAACTATTAATGAACAAAAAGACTAATGCTGATGCATTATCTAATAGT  840

Qy   841 TTTATTAAAAAGTGATTCAAACTACAGTTTGTTGTTGCTTTTAGTGTAACCCGTCAAT    900
Db   841 TTTATTAAAAAGTGATTCAAACTACAGTTTGTTGTTGCTTTTAGTGTAACCCGTCAAT    900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTTGTTGTTGATGTAACTACACCCGTCAAT       960
Db   901 AATGTTCAACCTTCAAACTACAGTTTTGTTGTTGATGTAACTACACCCGTCAAT       960

Qy   961 TATAAATATGCAAGAAGGACGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020
Db   961 TATAAATATGCAAGAAGGACGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020

Qy  1021 AACACGAATAGTATCACAGATGTTCTTGCCACTTCAACTAGTGAATAATAATACTACAGCTAATCCAACT 1080
Db  1021 AACACGAATAGTATCACAGATGTTCTTGCCACTTCAACTAGTGAATAATAATACTACAGCTAATCCAACT 1080

Qy  1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG   1140
Db  1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG   1140

Qy  1141 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAATTAAATAATGGAAAT    1200
Db  1141 TTGGTTAAAGCAGCTGATGCTGATAACGTTGGATTACAATACAATTAAATAATGGAAAT    1200

Qy  1201 GTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT   1260
Db  1201 GTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT   1260

Qy  1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT 1306
Db  1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT 1305
```

```
RESULT 3
I17387
LOCUS       I17387    1305 bp    DNA    linear    PAT 06-OCT-1996
DEFINITION  Sequence 1 from patent US 5489430.
```

```
                                    /organism="Mycoplasma gallisepticum"
                                    /mol_type="genomic DNA"
                                    /db_xref="taxon:2096"

ORIGIN

        Query Match           96.7%;   Score 1262.4;  DB 6;  Length 1387;
        Best Local Similarity 98.2%;   Pred. No. 4.4e-174;
        Matches 1283;  Conservative   0;  Mismatches  22;  Indels  1;  Gaps  1;

Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA  60
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAAACACAAAATCTTCTAACAA  60

Qy     61 AATCCTAATAAATAAGCCGTTAAATTAACTAAAAATTAAAAAATGGTTTTTTCTTATC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAATAAATAAGCCGTTAAATTAACTAAAAATTAAAAAATGGTTTTTTCTTATC  120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTTATTTTTATTTAGTGTCATCTTTAGAT  180
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTTATTTTTATTTTTACTCATCTTTAGAT  180

Qy    181 ATAAATATATTCTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATTCTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTG  240

Qy    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTTAGCTGTATGTCTATTACTAAAAA  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTTAGCTGTATGTCTATTACTAAAAA  300

Qy    301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCAGGCGGAATGGAGTAACAGAT  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGGAATGGAGTTAACTGAT  360

Qy    361 CTAATCAATGCTAAAGCGATGACATATGCCAAGATTGAAGCT  420
          |||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCGATGACATATGCCAAGATTGAAGCT  420

Qy    421 AGTTTATCATCTGCTTATAGTGAAGCTAAAAACTAATTGAACAATACCTAATGCAACATTA  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAAGCTAAAAACTAATTAACAATGCAACATTAACTA  480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT  540

Qy    541 AAAACGACTTTGATAATGAACACCAAATTAGTTGAAGCATCAAAGCACTAAAAACC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTGATAATGAACACCAAATTAGTTGAAGCATCAAAGCACTAAAAACC  600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT  660

Qy    661 CGCAATAATTTAGTGGATCTATACAAATAAAGCTAGTAGTTTAATAACTAAAACTAGAT  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTTAGTGGATCTATACAAATAAAGCTAGTAGTTTAATAACTAAAACTAGAT  720

Qy    721 CCACTAAATGGGGAACGCTTTAGATTCTAATGATTACTACAGTAATCGGAATATT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCACTAAATGGGGAACGCTTTAGATTCTAATGATTACTACAGTAATCGGAATATT  780

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840

Qy    841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTGTAGGGACTTTTGTGTCTGATGTAACAACGCT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTCTTTTGTTGTCTGATGTAACAACGCT  900

Qy    901 AATGTTCAACTTCAAACTACAGTTTTGTTGTCTGATGTAACACCCGTCAAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AATGTTCAACTTCAAACTACAGTTTTGTTGTCTGATGTAACACCCGTCAAT  960

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT  840

Qy    841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTCTTTTGTTGTCTGATGTAACAACGCT  900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTCTTTTGTTGTCTGATGTAACAACGCT  900

Qy    901 AATGTTCAACTTCAAACTACAGTTTTGTTGTCTGATGTAACACCCGTCAAT  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AATGTTCAACTTCAAACTACAGTTTTGTTGTCTGATGTAACACCCGTCAAT  960

Qy    961 TATAAATATGCAAGAGGACCGTTTGGAACTTGGATGGTGATGAACCTTCAAGTAGAAATTCTTGCA  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TATAAATATGCAAGAGGACCGTTNNNAATGGTGATGAACCTTCAAGTAGAAATTCTTGCA  1020

Qy   1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTAGCTGGAACAAACACGAAG  1080
          |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTTTCTNNNATTTATAGTTAGCTGGAACAAACACGAAG  1080

Qy   1081 TACCAATTTAGTTTTAGCAACTACGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACCAATTTAGTTTTAGCAACTACGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG  1140

Qy   1141 TTGGTTAAAGCAGCTGATTGGTGATTAAAATCGTTTTATCACAGTTAACATAATAATGGAAAT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGCTGATTGGTGATTACAATAACAAATTAATAATAATGGAAAT  1200

Qy   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAACTACAGTAATCCAACT  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAACTACAGTAATCCAACT  1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT  1306
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT  1305
```

```
RESULT 4
E09301
LOCUS        E09301                 1387 bp    DNA        linear      PAT 29-SEP-1997
DEFINITION   DNA encoding Mycoplasma TM-16.
ACCESSION    E09301
VERSION      E09301.1  GI:22025928
KEYWORDS     JP 195513295-A/2.
SOURCE       Mycoplasma gallisepticum
  ORGANISM   Mycoplasma gallisepticum
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE    1  (bases 1 to 1387)
  AUTHORS    Mori,H., Saito,S., Okawa,S., Funato,H., Iritani,K., Aoyama,S. and
             Takahashi,K.
  TITLE      NEW ANTIGEN PROTEIN, ITS GENE, RECOMBINANT BACULOVIRUS AND ITS USE
  JOURNAL    Patent: JP 195513295-A 2 23-MAY-1995;
             NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT      OS   Mycoplasma gallisepticum
             PN   JP 195513295-A/2
             PD   23-MAY-1995
             PF   27-AUG-1993 JP 1993213102
             PI   MORI HAJIME, SAITO SHUJI, OKAWA SETSUKO, FUNATO HIRONO, PI
             IRITANI KOICHI,
             PI   AOYAMA SHIGEMI, TAKAHASHI KIYOTO
             PC   C07K14/30,A61K39/00,A61K39/00,C12N7/01,C12N15/31,C12R21/02, PC
             (C12P21/02,
             PC   C12R1:92);
             CC   strandedness: Double;
             CC   topology: Linear;
             FH   Key            Location/Qualifiers
             FH
             FT   source         1..1387
             FT                  /organism='Mycoplasma gallisepticum'
             FT   CDS            202..1108
             FT                  /product='TM-16'.
FEATURES             Location/Qualifiers
     source          1..1387
```

```
Qy   961  TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020

Qy  1021  AACACGAATAGTATCACAGATGTTTCTTGGATTTAGTGTTTAGCTCGAACAAACACGAAG  1080
          ||||||||||||||||||||||||  ||||||||| |||||||||||||||||||||||||
Db  1021  AACACGAATAGTATCACAGATGTTTCTCTNNNAATTATAGTTTAGCTGGAACAAACACGAAG  1080

Qy  1081  TACCAATTAGTTTTAGCAACTGGATCCATCAACTGGTTATTTATATTTCCCTTATAAG  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TACCAATTAGTTTTAGCAACTGGATCCATCAACTGGTTATTTATATTTCCCTTATAAG  1140

Qy  1141  TTGGTTAAAGCAGCTGATGCTAATAAACGTTGGATTACAATACAAATTAAAATAATGGAAAT  1200
          |||||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||
Db  1141  TTGGTTAAAGCAGCTGATGCTAATAAACGTTGGATTACAATACAAATTAAAATAATGGAAAT  1200

Qy  1201  GTTCAACAAGTTGAGTTGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GTTCAACAAGTTGAGTTGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT  1260

Qy  1261  CCAGCAGTTGATGAGATTAAAAGTTGCTAAAAATCGTTTATCAGGTT  1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  -CAGCAGTTGATGAGATTAAAAGTTGCTAAAAATCGTTTATCAGGTT  1305
```

RESULT 5

```
LOCUS       AE016968            302070 bp    DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 2 of 4 of the complete
            genome.
ACCESSION   AE016968 AE015450
VERSION     AE016968.1  GI:31541264
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 302070)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
   PUBMED   12949158
REFERENCE   2  (bases 1 to 302070)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzell,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eaglevile Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..302070
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:233150"
     gene            79..3588
                     /locus_tag="MGA_1011"
                     /note="MGA_1011; MGR_224"
     CDS             79..3588
                     /locus_tag="MGA_1011"
                     /codon_start=1
                     /transl_table=4
                     /product="conserved hypothetical"
                     /protein_id="AAP56566.1"
                     /db_xref="GI:31541265"
                     /translation="MMNQNKKNNLPKKPLVKNKGIDLDDFLSEEIAVLDNFHLTKPKK
                     QQPLLLENKRAKKVSQPVVVKPGEKKVKAKKAKVVITDPRILKKRSRLLTIGFSLIGF
                     GITGPLLSFAILNQIDLSNQRTLHDLPIASFMIHNKLVRLLNQQLHFNYSWLYSDIAYLIRKTY
                     NKQSEINGIDLSNQRTLHDLPIASFMIHNKLVRLLNQQLHFNYSWIYSDIAYLIRKTY
                     DENKQAFIPLKTTDLQTKSSDNNPLHVYLTANFEFRNQTDSPQTVFVRYPNQRYEKQ
```

(continued, right column)

```
                     IQPKGILNLKLSLSASNDIDPLKPDYIGSIDPVDLAPGYINAVVTKDLYHNYYL
                     NWNLSLSELSMNDSNYRFSNFINNNVPSMNILLKKTIPGKDPYQAANDLVGENKLLN
                     RDLTPEQISRNLKFYFSNEFHSQIEWPFNLIRDVFKWLIVANNTNKKLYLVLSDNAAVF
                     AKLFSVNFIFKLNINYANSKNAILKLITDALSSNPIVNAKNVVEVLYDNFDFAFKLLIN
                     TYFFDVSAYSNLIDSLTKDVNSAFDPYQRLYELMPILRQLISNFPTKIKYFBIIBRIL
                     GYRDDAGNLIPIPRFNTYLINTIFESKRLLNDIYDQLLNISNVDFDMATANTNNGSES
                     WSRFTRLFNFLFIQNRSTLVNLFTQSTNYFDDLITLFISNNFNFLFQLLKLFGINVNT
                     LDLFFKTALDLIFLKNFEILSRDNLKAKINLFNWIIKSLDQTNINKIKFGLYKFENF
                     DNLVLTQSFTLKVKKLDLGFKLIDNGLQIPYGVLSALFEFLPNIKIVDFPIKYIVDDK
                     RYEBLKBKGYQQAINEGYSSLLNAGGIFSFKWAGDPAKNFPSSVFDDIWKNRILKGVS
                     DNNRTIKDLANEIFSGTNNNDPNRVIATINGAGYFKWSGENVDLLPFFEILPSNINQA
                     QIVYQLYNVDQSKDFSQIYQNLKLHFPRYSYRQPGIPPSDLSKKIYQSISGMVDQTLN
                     YSLNRKYLTKTRIILPGGDKKNHLILQENNPYENVALFYNANQSLINQYDRMNBQD
                     KRKLREQIMITDRRARVKGIDRDPKTRETSLLPQSLRMLDSLIQIKGVNPSLVIHTN
                     WSFSTISERISLNITTBVKFGLISKQVSINLAYETFLSAFSYQVLLPYPIIDLSNPNN
                     PVFKLKBEISHLVFSLELPGAGE"
     gene            3664..5019
                     /locus_tag="MGA_1014"
     CDS             3664..5019
                     /locus_tag="MGA_1014"
                     /note="MGA_1014; MGR_225"
                     /codon_start=1
                     /transl_table=4
                     /product="unique hypothetical"
                     /protein_id="AAP56567.1"
                     /db_xref="GI:31541266"
                     /translation="MTPYQKEIFDYDDEGFVDSIFHDDETDRFDEDQQDHFGFDSRTN
                     RIKRNVNYDQHDLQTLDVSRNKNRFFDKLVRSTKTYNFIAIWLNLFLITFVLAIVFF
                     YLNKFSPQQSESTYVSVNLFLTIALAAVSYFISVILYGFVKMHNKQIKYSLTEKSLDY
                     YSIFKLYVSYFLLGWIPLLGVIISVGSHIALQNLDDSDVSDNMLWYQESIDLLIDM
                     NGYENSYRELESKKVYLDQELQRLNYEKQLNAKGNBLAVVSSNALPNFKNSKQLLNY
                     VSNSGPLSASLVEQVKSAITQLREBRRKLRRENRKLRALLSKANNIDEQPLLITSSNN
                     LFNDGBQESLFNLTDQPSRKNEGLNLENFKDRIDWISENLNNDRAIVEIGIKKETAQ
                     NLRSRAPKVVKKNNQPTYKVKMVNNKKEAKVGRNQEBSDLTKQSNLLNFVNARIWVNQ
                     L"
     tRNA            5196..5269
                     /product="tRNA-Trp"
                     /note="codon recognized: UGG; tRNA genes predicted using
                     FASTRNA software of the Institut Pasteur, France;
                     functional prediction of the tRNAs is performed according
                     to Ushida C. and Muto"
     gene            6561..7661
                     /gene="hatA"
                     /locus_tag="MGA_1017"
     CDS             6561..7661
                     /gene="hatA"
                     /locus_tag="MGA_1017"
                     /note="MGA_1017; high affinity transport protein A (HatA)
                     [Mycoplasma gallisepticum]; MGR_226"
                     /codon_start=1
                     /transl_table=4
                     /product="HatA"
                     /protein_id="AAP56568.1"
                     /db_xref="GI:31541267"
                     /translation="MSKTIKKLFLISSSFLLSSAFVSSCAKANDANNQGTIQLKVDLK
                     IDHDANAIINDYLKNLSQTVSQKLNKPVSITSNFTBDGNTSIDRVKSNLTDLAFVSSS
                     AIKTNDTDYITKIQTLTRSFKFDNTPDFVEAGTLDQKARKISELFNEIPYRNWTDENR
                     MWDGNKYQFFYGPBDDLISFYRGMILIGGDQDTLAKIKBAWAQKNWATFSSYGIGGGR
                     TSSNGRYIYPLNLLKFHNDPNLTISDYQSVRGREIGTNPKVHIVFDDMNSFAWTQNK
                     NKNVPNYTFNPNDNNTKVEILSLTDPALNIGVFNKRLINDVINAVSEBSIVELAKNNQ
                     DTYGPTVGINGYRVIKDVNKEVYQPINKANIN"
     gene            7661..8407
                     /gene="hatB"
                     /locus_tag="MGA_1018"
     CDS             7661..8407
                     /gene="hatB"
                     /locus_tag="MGA_1018"
                     /note="MGA_1018; high affinity transport protein B (HatB)
                     [Mycoplasma gallisepticum]; MGR_227"
                     /codon_start=1
                     /transl_table=4
                     /product="HatB"
                     /protein_id="AAP56569.1"
```

```
gene        /db_xref="GI:31541268"
            /translation="MSNTNHIKFLDVSIKHGKRYKGFNVDHINPELKGSEILMLIGKSG
            SGKTTLLNSITDKLVSQGEIVFNDQNLPSLSCHKIRKIAKFIGFLDQIPNLILDDYV
            YENIIRSIRKPWYYKFNITPISVLNKVEBVLKTDLSDHINKLIKDLSGGQQRVDI
            AKIFFQKPKLLIDEPTTGLDYLNSEIVIREIKKLSEQFKAPVIISIHDLEIVKKFAD
            KILIIQNKKPILINNSADLNIDEIRKKYET"
            8379..9998
            /gene="hatC"
            /locus_tag="MGA_1019"

CDS         8379..9998
            /gene="hatC"
            /locus_tag="MGA_1019"
            /note="MGA_1019; high affinity transport protein C (HatC)
            [Mycoplasma gallisepticum]; an alternative start codon may
            be 27 nucleotides upstream: 'tta aca att ctg cgg atc tta
            ata tcg ATG' / 'LTILRILISM'; MGR_228"
            /codon_start=1
            /transl_table=4
            /product="HatC"
            /protein_id="AAP56570.1"
            /db_xref="GI:31541269"
            /translation="MRLEKNMRHNKFFQLLKINKFTVFTVFLFLCIISAIVIKLGIN
            AHGLAVFKNNISNFNQSAVNKSSLLTKSFEQLWITIKYTATGTPIGFLLGFLLGYLS
            SLRITNNFTAIFVKLGIIFPRSFPVVIFNLPNTSPNSDLSAIIIISWFSMLWNAKYI
            ADYIENSKVEQFKKFVNRSQNHLSSFINNIFIDIKSKIIVLFAYSLESNFRWTTILSA
            VGLIGIGQLINDPISINNDFASTLPLLVMVFLLINEGALYLFPETVLVARKSYPKTK
            NLNKLLTIKKYLVLVILFILIITFSIYSISFSIFSLLRLTIILFTDFFQRLFNFNKSFFAT
            TTFNENPLLMIILLTLQAILILGIVFFISLLFATLCSNLLNRYVSLFFKALFLIRTI
            PLIIVFRLFNPLFNSGISTVFIGSLYFSTSLAKKIYVLINSINWTVVTSLKSKLYTN
            FEILRVYIIPSIRKDLVTVYQLEFPSILRTIIILGAYGTSVIGQLLDIYIYRGSIENL
            GSYVISIMVYFQVIDLLSIMVRYKKFFYTNK"
            10143..11567
            /gene="proS"

gene        /locus_tag="MGA_1022"
            10143..11567
            /gene="proS"

CDS         /locus_tag="MGA_1022"
            /note="MGA_1022; [J] COG0442 Prolyl-tRNA synthetase;
            MGR_229"
            /codon_start=1
            /transl_table=4
            /product="ProS"
            /protein_id="AAP56571.1"
            /db_xref="GI:31541270"
            /translation="MAFFKKNLASSKDDNISDWYTEVISKGNLIHYSNIKGFMSFLPNG
            WKIWQLIKKELTKEFQRDILELCLSLISQNDFPLLEKKHLEGPAPELPIVTKISNET
            EKYILRPTSIAFCNLWRETLRNYRQLPIKHQWTSVFRVEKNTRPFLRNSEFFWHEI
            HSCFETEBEQSNDFAVDIWKLYQRFIKDILCIPLVAGEKTEAEKFAGAKTTYTVETIMP
            DGQALQSATSRNLSQNFSKAPDIRYQTKNNDYQNVFSMSAGVSTRIIGAIIMTHGDDD
            GLVFPTKVAPYHISLNCIFPDTNQELNAKLKELANKYSQKYRVHLNVNKDSTGEIIKN
            SQLRGDCCVLLMGPNDLKKNEIVPIDRITKQKQFINLDHLDQKLBELFSTFDQKLYQK
            AKAVFETKVDPAQTFEEFEQKIASGKFVRVFYCNEDLYEKQIKEKTGASSRCIIKYLD
            EQTQ8RCFISNKKAKVEIYFARSY"
            11768..12190
            /locus_tag="MGA_1027"

gene        11768..12190
            /locus_tag="MGA_1027"

CDS         /locus_tag="MGA_1027; MGR_230"
            /codon_start=1
            /transl_table=4
            /product="unique hypothetical"
            /protein_id="AAP56572.1"
            /db_xref="GI:31541271"
            /translation="MKMFKVKIAKKIFHTFVKKNHILNRQPERSISMNRKIINDAKKY
            LERYDPYRKRIENRQIIKSVFNHDIGVDNQSEMQSYVDLIDATLKRLKDEKDILIDV
            FLKRKHRNEFNFYSQASFYRLLNKACCSLFYSLGVVIPN"
            12194..13276
            /locus_tag="MGA_1029"

gene        12194..13276
            /locus_tag="MGA_1029"

CDS         /locus_tag="MGA_1029; MGR_231"
            /codon_start=1
            /transl_table=4
```

```
                /product="unique hypothetical"
                /protein_id="AAP56573.1"
                /db_xref="GI:31541272"
                /translation="MLISYNIDSNNNSKPIVINKEGWWWRLNPLRNKDGTIVVDKNSN
                PIYFRALLSFQNTWSDQNNLNLYFEILFENNSILPPNTIQSNVFIAGNVPSVFSEIT
                NEETDGYQKAQIKVSVPLEKLFTKERFEEGIDISVVIDLYLQNDILSGIRDSLESNALR
                LLLENLKDARLVTKIKLNNFLPFRDRVGNRDDADYQVQEKKAISWYQPNYSFLSNLD
                QIRDLLTINVDAVTLFKSRVVFDYSTTNNGSTKSLVIENTNQITYPPNTDIKLALDT


    Query Match           90.3%;   Score 1179.4;   DB 1;   Length 302070;
    Best Local Similarity 95.3%;   Pred. No. 8.4e-163;
    Matches 1249;   Conservative 0;   Mismatches 56;   Indels 6;   Gaps 3;

    Qy       1 AAAAACATCAGATGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA     60
               ||||||||||| ||||||||||||||| |||||| ||||||||| |||||||||||||||
    Db  189913 AGAACAAATCAGATTGTTAATCTGATATCTGATATCTTTGCTT-AAAAAACACAAAATCTTTTAACAA 189971

    Qy      61 AATCCTAAATAAATAAAGCCGTAAATAACTAAAAAAATTAAAAAAATTAAAAAAATGTTTTTCTTATC    120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db  189972 AATCCTAAATAAATAAAGCCGTAAATAACTAAAAAAATTAAAAAAATTAAAAAAATGTTTTTCTTATC 190031

    Qy     121 AACCAAAATTCTCTAGTAATAAACGCTTATTTAT-TTTTATTTTTAGTCATCTTTTAAGA    179
               ||||||||||| |||||||||||||||||||||| |||||||||||||||||| |||||||
    Db  190032 AACCAAAATTCTTTAGTAATAAACGCTTATTTATTTATTTTTTTTTAGTCATCATTTAAGA 190091

    Qy     180 TA----TAAATATATCTTAATATCTCATGAATAAGAAAAGAATCATCTTAAAGACTATTA    235
               ||    ||||||||||||||||||| |||||| |||||||||||||||||||||||||||
    Db  190092 TATATCTATATATATCTTAATATTCTATGAGTAAAAAAAGAATCATCTTAAAGACTATTA 190151

    Qy     236 GTTTGTTAGGTACACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTA    295
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db  190152 GTTTGTTAGGTACACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTA 190211

    Qy     296 AAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTGAAGCAGCGCGAATGGAGTTAA    355
               ||||||| |||| ||||||||||| ||||| ||||||||| ||||||||||||||||||
    Db  190212 AAAAAGACGCAAATCAAATAATAGCCAAATCAAACCCAATTACAACAGCGCGAATGGAGTTAA 190271

    Qy     356 CAGATCTAATCAATGCTAAAGCTAAAGCAATGACAATGACATTAGCTTCACTACAGACATATGCCAAGATTG    415
               | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db  190272 CTGATCTAATCAATGCTAAAGCTAAAGCAATGACAATGACATTAGCTTCACTACAGACATATGCCAAGATTG 190331

    Qy     416 AAGCTAGTTTATCATCTGCTTATAGTGAAGCTAATTAGAATCAGCCATCAACCAAGCTAATGCAA    475
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    Db  190332 AAGCTAGTTTATCATCTGCTTATAGTGAAGCTAATTAGAATCAGCCATCAACCAAGCTAATGCAA 190391

    Qy     476 CATTAGAACAACTAAAAATGGCTAAAACTAAATTAGAATCAGCACCCAAATTAGTTGAAGCATACAAAGACTAA    535
               || ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||| ||||||
    Db  190392 CACTAGAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCACCCAAATTAGTTGAAGCATACAAAACACTAA 190451

    Qy     536 CGGATAAAAGACTTTGATAATGAACACACCCAAATTAGTTGAAGCATACAAAGACTAA    595
               ||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||
    Db  190452 CGGATAAAAACCACTTTGATAATGAACACACCCAAATTAGTTGAAGCATACAAAACACTAA 190511

    Qy     596 AAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATC    655
               |||||||||| | ||||||||||||||||||||||||||| ||||||||||||||| |||
    Db  190512 AAACCACTTTATCACAACGTGCTACTAACCTTGAAGGTTTAGCATCAACTGCTTATATATC 190571

    Qy     656 AAATTCGCCAATAATTTAGTGGATCTATACATAATTAGTTAATAACTAAAAACAC    715
               | |||||| |||||||||||||||||||||||||||||||||||||||||||||
    Db  190572 AGATTCGTAATAATTTAGTGGATCTATACATAATTAGTTAATAACTAAAAACAC 190631

    Qy     716 TAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGA    775
               ||||||||||||||||||||||||||||||||||||||||||||||||| ||||| |||
    Db  190632 TAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGCTAATAAGA 190691

    Qy     776 ATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTAATGCATTACTA    835
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
    Db  190692 ATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTAATGCATTAGCTA 190751

    Qy     836 ATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTAGGGACTTTTACAA    895
               |||||||||| ||||| |||||||||||||| |||||||||||||| || ||||||||
    Db  190752 ATAGTTTTATCAAAGAAGTGATTCAAAATAATAAACAAAGTTTGTAGGAATGTTTACAA 190811
```

Qy   896 ACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGCTTTAGTGCTGATGTAACACCCG 955
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190812 ACACTAATGTTCAACCTTCAAACTATCAGTTTGTTGCTTTAGTGCTGATGTAACACCTG 190871

Qy   956 TCAATTATAAATATGCAAGAGAAGAGACCGTTTGAATGGTGTGATGAACCTTCAAGTAGAATTC 1015
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190872 TTAATTATAAATATGCAAGAGAAGAGACCGTTTGAAATGGTGATGAACCTTCAAGTAGAATTC 190931

Qy 1016 TTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTATAGTTTAGCTGGAACAAACA 1075
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190932 TTGCAAACACGAATAGTATCACAGATGTTTCTTGAATTTATAGTTTAGCTGGAACAAACA 190991

Qy 1076 CGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTT 1135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 190992 CGAAGTACCAATTTAGTTTTAGCAACTACGGTCCATCAACTGGTTATTTATATTTCCCTT 191051

Qy 1136 ATAAGTTGGTTAAAGCAGCTGTAATAACGTTGGATTACAAATTAGTTGGATTCTATACAAATAATG 1195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191052 ATAAGTTCGTTAAAGCAGCAGCCGTTTGGATTAGCTGTAATGTTGGATTCTATACAAATAATG 191111

Qy 1196 GAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTCAGTCGCAAATAATAATCACAGCTAATC 1255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191112 GAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTCAGTCGCAAATAATAATCACAGCTAATC 191171

Qy 1256 CAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 191172 CAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 191222

RESULT 6
AR035278
LOCUS   AR035278   2014 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5871742.
ACCESSION   AR035278
VERSION   AR035278.1 GI:5951946
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2014)
  AUTHORS   Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y.,
            Aoyama,S. and Takahashi,K.
  TITLE   Recombinant Avipox virus encoding polypeptide of mycoplasma
            gallisepticum, and utilized a live vaccine
  JOURNAL   Patent: US 5871742-A 7 16-FEB-1999;
FEATURES   Location/Qualifiers
     source   1..2014
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   83.2%; Score 1086.8; DB 6; Length 2014;
Best Local Similarity   96.7%; Pred. No. 1.2e-148;
Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

Qy   150 TTTATTTTTATTTTTAGT-CATCTTTAAGATATAAATATAATCTTAATATTCTATGAATA 208
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1 TTTATTTTTATTTTTGGTAAATCTTTAAGATATAAATATAAATATTTAATATTCTATGAATA   60

Qy   209 AGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTG 268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAAAAGAATCATCTTAAAGACTATTAGCTTGTTAGGTACAACATCCTTTCTTAGTATTG 120

Qy   269 GGATTTCTAGCTGTATGCTGATGTCTAATACTAAAAAGAATGCAAACCCAAATAGCCAACCC 328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGATTTCTAGCTGTATGCTGATGTCTAATACTAAAAAGAATGCAAACCCAAATAGCCAACCC 180

Qy   329 AATTAGAAGCAGCGCGAATGAGTGAGTCAACAGATCTAATCAATGCTAAAGCGGATGACATTAG 388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATTAGAAGCAGCGCGAATGGGAGTGAGTCAACAGATCTAATCAATGCTAAAGCGGATGACATTAG 240

Qy   389 CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG 448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   241 CTTCACTACAAGACTATGCCAAGATTGAAGCATGCAACATTGAAGCAACTAAAAAATGGCTAAACTAATT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   449 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAAATGGCTAAACTAATT 508

Db   301 AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAAATGGCTAAACTAATT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   509 TAGAATCAGCCATCAACCAAGCTAATACGATAAAACGACTTTGATAATGAACACCCAA 568

Db   361 TAGAATCAGCCATCAACCAAGCTAATACGATAAAACGACTTTGATAATGAACACCCAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   569 ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTAGAACAACACGTACTAACCTTG 628

Db   421 ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTAGAACAACACGTGCTACTAACCTTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   629 AAGGTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATA 688

Db   481 AAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   689 AAGCTAGTAGTTTAATAACTACTAAAACACTAGATCCAATCAATAATAATCAAACTAAATAATG 748

Db   541 AAGCTAGTAGTTTAATAACTACTAAAACACTAGATCCAATCAATAATAATCAAACTAAATAATG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   749 CTAATGAGATTACTACAGCTAATCGGAATATTAATAATAATCGTTATCAACTATTAATGAAC 808

Db   601 CTAATGAGATTACTACAGCTAATAGAATATTAATAATAATCGTTATCAACTATTAATGAAC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   809 AAAAGACTAATGCTGATGCATTACTCAATAATAGTTTTATTTATAAAAGTGATTCAAAATAATG 868

Db   661 AAAAGACTAATGCTGATGCATTAGTGCATTAGTTTTATTTATTAAAAGAGTGATTCAAAATAATA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   869 AACAAAGTTTGTAGGGACTTTACAAACGCTAATGTTCAACCTTCAACCTCAACTACAGTTTTG 928

Db   721 AACAAAGTTTTGTAGGAATGTTTACAAACACTAATGTTCAACCTTCAACCTCAACTACATAGTTTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   929 TTGCTTTAGTGCTGATGTAACACCCGTCAATTATAAAATGCAAGAAGGACCGTTTGGA 988

Db   781 TTGCTTTAGTGCTGATGTAACACCTGTTAATTATAAAATGCAAGAAGGACCGTTTGAA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   989 ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTT 1048

Db   841 ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATTACTACTGATGTTTCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1049 GGATTTATAGTTTAGCTGGAACAAACACGAAGTACCAATTTAGTTTTTAGCAACTATGGTC 1108

Db   901 GAATTTATAGTTTATCTGGAACAAACACGAAATATACCAATTACCAATTTAGTTTTTAGCAACTACGGTC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1109 CATCAACTGGTTATTATATTTCCCTTATATATAGTTGGTTAAAGCAGCTGGTCAATAACG 1168

Db   961 CATCAACTGGTTATTTATATTTCCCTTATATATATGGTTAAAGCAGCTGATGCTAGTAATG 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1169 TTGGATTACAATACAAATAAAATAAATAAAAGTTGAAATGTTCAACAAGTTGCCACTTCAA 1228

Db 1021 TTGGATTACAATACAAACTAAATAAACTAAATAAAAGTTGAAATGTTCAACCAGTTGCCACTTCAA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1229 CTAGTGCAAATAATATACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA 1288

Db 1081 CTAGCGCAAATAAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1289 AAATCGTTTTATCAGGTT 1306

Db 1141 AAATCGTTTTATCAGGTT 1158

RESULT 7
AX665164
LOCUS   AX665164   1152 bp   DNA   linear   PAT 26-MAR-2003
DEFINITION Sequence 1 from Patent EP1275716.
ACCESSION   AX665164
VERSION   AX665164.1 GI:29290294
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum
  ORGANISM Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

REFERENCE   1
AUTHORS   Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki.Y.
TITLE   Modified dna molecule, recombinant containing the same thing, and
            uses thereof
JOURNAL   Patent: EP 1275716-A 1 15-JAN-2003;
            Zeon Corporation (JP)
FEATURES   Location/Qualifiers
     source   1..1152
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:2096"
                     /note="TM-1 gene (after EcoRI)"
ORIGIN

Query Match   78.7%; Score 1028; DB 6; Length 1152;
Best Local Similarity 100.0%; Pred. No. 4.8e-140;
Matches 1028; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    279 CTGTATGTCTATTACTAATAAAAGATGCAATAAATGGCCAAACCCAATTAGAAGC 338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 CTGTATGTCTATTACTAATAAAAGATGCAATAAATGGCCAAACCCAATTAGAAGC 65

Qy    339 AGGCGGAATGGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 AGCCGGAATGGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 125

Qy    399 AGACTATGCCAAGATTGAAGCTAGTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAA 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 AGACTATGCCAAGATTGAAGCTAGTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAA 185

Qy    459 CAATAACCTAATGCAACATTAGAACACAACTAAAAATGGCTAAAACTAATTAGAATCAGC 518
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 CAATAACCTAATGCAACATTAGAACACAACTAAAAATGGCTAAAACTAATTAGAATCAGC 245

Qy    519 CATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTAGTTGA 578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 CATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTAGTTGA 305

Qy    579 AGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAGGTTTGTC 638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 AGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAGGTTTGTC 365

Qy    639 ATCAACTGCTTATAATCAATTGCAATAATTAGTGGATCTATACAATAAAGCTAGTAG 698
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 ATCAACTGCTTATAATCAATTGCAATAATTAGTGGATCTATACAATAAAGCTAGTAG 425

Qy    699 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCGTTTAGATTCTAATGATGAT 758
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCGTTTAGATTCTAATGATGAT 485

Qy    759 TACTACAGTTAATCGGAATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 818
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 TACTACAGTTAATCGGAATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 545

Qy    819 TGCTGATGCATTATCATAGTAGTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT 878
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 TGCTGATGCATTATCATAGTAGTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTT 605

Qy    879 TGTAGGGACTTTACAAACGTCTAATGTTCAACCTTCAACTACAGTTTTGTTGCTTTTAG 938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 TGTAGGGACTTTACAAACGTCTAATGTTCAACCTTCAACTACAGTTTTGTTGCTTTTAG 665

Qy    939 TGCTGATGTAACACCCGTTAACAGGACCGTTGGAATGGTGATGA 998
          ||||||||||||||||||||||||||||||||||||||||||||
Db    666 TGCTGATGTAACACCCGTTAACAGGACCGTTGGAATGGTGATGA 725

Qy    999 ACTTTCAAGTAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTTTGGATTTATAG 1058
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 ACTTTCAAGTAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTTTGGATTTATAG 785

Qy   1059 TTTAGCTGGAACAAACAGAGTACCAATTTAGTTTAGCAACTATGTCCATCAACTGG 1118
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 TTTAGCTGGAACAAACAGAGTACCAATTTAGTTTAGCAACTATGTCCATCAACTGG 845

Qy   1119 TTATTTATATTTCCCTTATAGAAGTTGGTTAAAGCAGCTAATAACGTTGGATTACA 1178
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 TTATTTATATTTCCCTTATAGAAGTTGGTTAAAGCAGCTAATAACGTTGGATTACA 905

Qy   1179 ATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 1238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 ATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 965

Qy   1239 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT 1298
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTT 1025

Qy   1299 ATCAGGTT 1306
          ||||||||
Db   1026 ATCAGGTT 1033
```

RESULT 8
E09896
LOCUS   E09896   1015 bp   DNA   linear   PAT 29-SEP-1997
DEFINITION Mycoplasma gallisepticum TM-1 gene.
ACCESSION   E09896
VERSION   E09896.1 GI:22026524
KEYWORDS   JP 199236498-A/2.
SOURCE   Mycoplasma gallisepticum
  ORGANISM Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 1015)
  AUTHORS   Saito,S., Kyo,T., Okawa,S. and Iritani,K.
  TITLE   DETECTION OF MYCOPLASMA GALLISEPTICUM AND NEW DNA TO BE USED
  JOURNAL   Patent: JP 199236498-A 2 12-SEP-1995;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
OS   Mycoplasma gallisepticum
PN   JP 199236498-A/2
PD   12-SEP-1995
PF   25-FEB-1994 JP 194052764
PI   SAITO SHUJI, KYO TSUGUO, OKAWA SETSUKO, IRITANI KOICHI PC
            C12Q1/68,C12N15/09,(C12Q1/68,C12R1:35),(C12N15/09,C12R1:35); CC
            strandedness: Double;
CC   topology: Linear;
FH   Key   Location/Qualifiers
FH
FT   source   1..1015
FT   /organism='Mycoplasma gallisepticum'   FT CDS
FT   202..987
FT   /product='TM-1 gene product'
FT   misc_feature   718..741
FT   /note='this region is favorble for   FT
                     hybridization probe to
                     detect TM-1 gene'.
FT   Location/Qualifiers
     source   1..1015
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
ORIGIN

Query Match   75.8%; Score 989.4; DB 6; Length 1015;
Best Local Similarity 98.4%; Pred. No. 2e-134;
Matches 999; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAACACAAAATCTTCTAACAA 60
          ||||||||||||||| |||||||||||||||||||||||    |||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAAATAAATAAGCCGTTAAATAACTAACTAAAATTAAAAAATGGTTTTCTTATC 120
          ||||||||||||||||||||||||||||||      |||||||||||||||||||||||
Db     61 AATCCTAAAATAAATAAGCCGTTAAATAACGAATAGTATCAGATAAAAAATGGTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTTAAGAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTTAAGAT 180
```

JOURNAL   Mycoplasma gallisepticum as a possible protective antigen
           Vaccine 11 (10), 1061-1066 (1993)
MEDLINE   94025893
PUBMED   8212828
REMARK   GenBank staff at the National Library of Medicine created this
           entry [NCBI gibbsq 138145] from the original journal article.
           This sequence comes from Fig. 2.
FEATURES   Location/Qualifiers
     source   1..1015
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
     gene   202..>1015
                     /gene="TM-1"
     CDS   202..>1015
                     /gene="TM-1"
                     /note="29 kda polypeptide; This sequence comes from Fig.
                     2; conceptual translation presented here differs from
                     translation in publication"
                     /codon_start=1
                     /transl_table=4
                     /product="TM-1"
                     /protein_id="AAB28343.2"
                     /db_xref="GI:7249262"
                     /translation="MNKKRIILKTISLLGTTSFLSIGISSCMSITKDANPNNGQTQL
                     QAARMELTDLINAKARTLASLQDYAKIEASLSSAYSEAETVNNNLNATLEQLKMAKTN
                     LESAINQANTDKTTFDNEHPNLVEAYKALKTLEQRATNLEGLASTAYNQIRNNLVDL
                     YNNASSLITKTLDPLNGGMLLDSNETTVNRNINNLSTINEQKTNADALSNSFIKKV
                     IQNNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRI"

ORIGIN

Query Match   75.5%;   Score 986.2;   DB 1;   Length 1015;
Best Local Similarity   98.2%;   Pred. No. 5.9e-134;
Matches 997;   Conservative   0;   Mismatches 18;   Indels   0;   Gaps   0;

Qy   1 AAAAACATCAGATTGTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 AAAAACATCAGATTGTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy   61 AATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTCTTATC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAAATTAAAAAAAATGGTTTTCTTATC 120

Qy   121 AACCAAAATCTCTAGTAATAAACGCTTATTATTTTATTTTTAGTCATCTTTTAAGAT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AACCAAAATCTCTAGTAATAAACGCTTATTATTTTATTTTTAGTCATCTTTTAAGAT 180

Qy   181 ATAAATATATCTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ATAAATATATCTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy   241 TTAGGTACAACATCCTTCTTAGCATTGGAATTGGGATTTCCAGCTGTATGTCATTACTAAAAAA 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTAGGTACAACATCCTTCTTAGCATTGGAATTGGGATTTCCAGCTGTATGTCATTACTAAAAAA 300

Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGAAGCT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GACGGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGAATGAAGCT 360

Qy   361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCTAAGATTGAAGCT 420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy   481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy   541 AAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAGCACTAAAAACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAGCACTAAAAACC 600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTGCTACTAACTGCTTATAATCAGATT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTGCTACTAACTGCTTATAATCAGATT 660

Qy   661 CGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGAT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CGTAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAAACACTAGAT 720

Qy   721 CCACTAAATGGGGAACGCTTTAGATGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCACTAAATGGGGAACGCTTTAGATGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy   781 AATAATACGTTATCAACATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AATAATACGTTATCAACATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy   841 TTTATTAAAAAGTGATTCAAAATAAAAATAATGAACAAAGTTTTGTAGGGACTTTTACAAACGCT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TTTATTAAAAAGTGATTCAAAATAAAAATAATGAACAAAGTTTTGTAGGGACTTTTACAAACGCT 900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTTGTTTGCTTTTAGTGTCGTGATGTAACACCCGTCAAT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 AATGTTCAACCTTCAAACTACAGTTTTGTTTGCTTTTAGTGTCGTGATGTAACACCCGTCAAT 960

Qy   961 TATAAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TATAAAATATGCAAGAAGGACCGTTTGAAATGTGTGATGAACCTTCAAGTAGAATTC 1015

RESULT 9
S65869
LOCUS   S65869   1015 bp   DNA   linear   BCT 16-MAR-2000
DEFINITION TM-1=29 kda polypeptide [Mycoplasma gallisepticum, S6, ATCC 15302]
            Genomic, 1015 nt].
ACCESSION   S65869
VERSION   S65869.1   GI:425376
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum
  ORGANISM Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 1015)
  AUTHORS   Saito,S., Fujisawa,A., Ohkawa,S., Nishimura,N., Abe,T., Kodama,K.,
            Kamogawa,K., Fujisawa,A., Aoyama,S., Iritani,Y. and Hayashi,Y.
  TITLE   Cloning and DNA sequence of a 29 kilodalton polypeptide gene of

```
Db    181 ACGGATAAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 AAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTA 600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATATAATCAAATT 660
           |||||||||||||||||||||||||||||||||||||| |||||||||||
Db    595 AAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATAAT 654

Db    241 AAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATAAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTAGCTTCAACTGCTTATAATCAGATT 660

Qy    661 CGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTAATAACTAGACACTAGAT 720
           |||||||||| ||||||||||||||||||||||||||||||||||||||
Db    655 CAAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTAATAACTAAAACA 714

Db    661 CGTAATAATTAGTGGATCTATACAATAATGCTAGTAGTTAATAACTAGACACTAGAT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    301 CAAATTCGCAATAATTAGTGGATCTATACAATAATGCTAGTAGTTAATAACTAAAACA 360

Qy    721 CCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
           |||||||||||||||||||||| |||||||||||||||||||||||||||
Db    715 CTAGATCCACTAAATGGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGG 774

Db    721 CCACTAAATGGGGGAACGGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361 CTAGATCCACTAAATGGGGGAACGGTTTTAGATTCTAATGAGATTCTAATCGGCTAATAAG 420

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840
           |||||||||||||||||||||||||||| |||||||||||||||||||
Db    775 AATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCT 834

Db    781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421 AATATCCAGAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCT 480

Qy    841 TTTATTAAAAAAGTGATTCAAACTACAGTTTGTAGGGACTTTTACAAACGCT 900
           |||||||||||||||||||||||||| ||||||||||||||
Db    835 AATAGTTTATTAAAAAAGTGATTCAAATAATGAACAAAGTTTGTAGGGACTTTTACA 894

Db    841 TTTATTAAAAAAGTGATTCAAACTACAGTTTGTAGGGACTTTTACAAACGCT 900
           ||||||||||||||||||||||||||||||||||||||||||||
Qy    481 AATAGTTTATTAAAAAAGTGATTCAAATAATGAACAAAGTTTGTAGGGACTTTTACA 540

Qy    901 AATGTTCAACTTCAAACTACAGTTTTGTTGCTTTAGTGCTGATGTAACACCCCGTCAAT 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895 AACGCTAATGTTCAACTTCAAACTACAGTTTTGTTGCTTTAGTGCTGATGTAACACCC 954

Db    901 AATGTTCAACTTCAAACTACAGTTTTGTTGCTTTAGTGCTGATGTAACCCGTCAAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||
Qy    541 AACGCTAATGTTCAACTTCAAACTACAGTTTTGTTGCTTTAGTGCTGATGTAACACCC 600

Qy    961 TATAAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT 1015
           ||||||||| ||  |||||||||||||||||||||||||||||||||||
Db    955 GTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT 1014

Db    961 TATAAATATGCAAGAACGGTTTGAAATGGTGATGAACCTTCAAGTAGAATTC 1015
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    601 GTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATT 660

Qy   1015 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGTGGGACAAAC 1074
           |||||||||||||||||||||||||||||||||||| |||||||||||||
Db    661 CTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGTCGGAACAAAC 720

Qy   1075 ACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCT 1134
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCT 780

Qy   1135 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACATTACAATACAAATTAATAAT 1194
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACATTACAATACAAATTAATAAT 840

Qy   1195 GGAAAGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAAT 1254
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GGAAAGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCCACAGCAGCAGACTACAGCTAAT 900

Qy   1255 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 952
```

```
RESULT 11
E02348
LOCUS       E02348                853 bp      DNA         linear    PAT 29-SEP-1997
DEFINITION  DNA sequence coding for TMG-1.
ACCESSION   E02348
VERSION     E02348.1  GI:2170583
KEYWORDS    JP 199011795-A/7.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   Kodama,K., Saito,S., Yanagida,N., Kamogawa,K., Iritani,K. and
            Aoyama,S.
  TITLE     FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE
            THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME
  JOURNAL   Patent: JP 199011795-A 7 24-APR-1990;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
COMMENT     OS    Mycoplasma gallisepticum
            PN    JP 199011795-A/7
            PD    24-APR-1990
            PF    02-JUN-1989 JP 198914 0283
```

```
RESULT 10
AX665187
LOCUS       AX665187              1082 bp     DNA        linear     PAT 26-MAR-2003
DEFINITION  Sequence 24 from Patent EP1275716.
ACCESSION   AX665187
VERSION     AX665187.1  GI:29290312
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1
  AUTHORS   Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
  TITLE     Modified dna molecule, recombinant containing the same thing, and
            uses thereof
  JOURNAL   Patent: EP 1275716-A 24 15-JAN-2003;
            Zeon Corporation (JP)
FEATURES             Location/Qualifiers
     source          1..1082
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:2096"
                     /note="Modified TTM-1 portion (downstream of BglI) of
                     pNZ40K-S"
ORIGIN

Query Match     70.6%;  Score 921.6;  DB 6;  Length 1082;
Best Local Similarity   98.0%;  Pred. No. 1.4e-124;
Matches   933;  Conservative    0;  Mismatches    19;  Indels    0;  Gaps    0;
```

```
Qy    355 ACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 414
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGACTATGCCAAGATT 60

Qy    415 GAAGCTAGTTATCATCGTATTAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA 474
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAAGCTAGTTATCATCGTATTAGTGAAGCTGAAACAGTTAACAATAACCTGCAGGCA 120

Qy    475 ACATTAGAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCACCAAGCTAAT 534
           |||||||||||||||||||||||||||||||| ||||||||||||||||
Db    121 ACATTAGAACAACTAAAAATGGCTAAAAATAATTAGAAATCAGCCATCACCAACCAAGCTAAT 180

Qy    535 ACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 594
```

```
PR   02-JUN-1988 JP 88P   136343
PI   KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,
PI   IRITANI KOICHI, AOYAMA SHIGEMI
PC   C07K13/00,C07K7/06,C07K15/04,C12N1/21,C12N15/31//A61K37/02, PC
     C12P21/02,
PC
G01N33/569,(C12N1/21,C12R1:19),(C12N15/31,C12R1:35),(C12P21/02, PC
C12R1:19];
CC   strandedness: Double;
CC   topology: Linear;
CC   hypothetical: No;
CC   anti-sense: No;
CC   *source: strain=S6;
FH   Key                Location/Qualifiers
FT   5'UTR              1..39
FT   CDS                40..825
FT                      /product='TMG-1'
FT   3'UTR              826..853.

FEATURES             Location/Qualifiers
     source          1..853
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
ORIGIN

Query Match             63.1%;  Score 824.2;  DB 6;  Length 853;
Best Local Similarity   97.9%;  Pred. No. 2e-110;
Matches   835;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy    163 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATC 222
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db      1 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATC  60

Qy    223 TTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTCTTAGCATTCTAGCTGT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTCTTAGCATTCTAGCTGT 120

Qy    283 ATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCG 342
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    121 ATGTCTATTACTAAAAAAGACGCAAACCCAAATAATGGCCAAACCCAATTACCAAGCAGCG 180

Qy    343 CGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTACCACTACAAGAC 402
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGAATGGAGTTAACTGATCTAATCAATGCTAAAGCCGATGACATTACCACTACAAGAC 240

Qy    403 TATGCCAAGATTGAAGCTAGTTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 462
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATGCTAAGATTGAAGCTAGTTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 300

Qy    463 AACCTTAATGCAACATTAGACAACAACTAAAAATGGCTAAAAACTAATTAGAATCAGCAGCCATC 522
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AACCTTAATGCAACACTAGACAACAACTAAAAATGGCTAAAAACTAATTAGAATCAGCAGCCATC 360

Qy    523 AACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCA 582
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    361 AACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCA 420

Qy    583 TACAAAGCACTAAAAACCACTTTAGACAACAACCGTACTAACCGTGTTGTCATCA 642
          ||||||||||||||||||||||||||||||||||||||||||||
Db    421 TACAAAGCACTAAAAACCACTTTAGACAACAACCGTACTAACCGTGAAGGTTTAGCTTCA 480

Qy    643 ACTGCTTATAATCAATTCGCAATAATTCGGATCTATACAATAAGCTAGTAGTTTA 702
          ||||||||||||||| |||| |||||||||| ||||||||||||| ||||||||||
Db    481 ACTGCTTATAATCAGATTCGTAATAATTCGTGGATCTATACAATAATGCTAGTAGTTTA 540

Qy    703 ATAACTAAAAACACTAGATCCACTAAAATGGGGAACGCTTTTAGAATTCTAATGAGATTACT 762
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db    541 ATAACTAAAAACACTAGATCCACTAAAATGGGGAATGCTTTTAGAATTCTAATGAGATTACT 600

Qy    763 ACAGTTAATCGGAATATTAATTATTAATAATACGTTATCAACTATTAATCAACATTAATGAACAAAAGACTAATGCT 822
          |||||||||||||||   |||||||||   |||||||||||   ||||||||||||||||||||||||||||||
Db    601 ACAGTTAATCGGAATTTATCAACTATTAATAATACGTTATCAACTATTAATCAACATTAATGAACAAAAGACTAATGCT 660
```

```
Qy    823 GATGCATTATTCTAATAGTTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTA 882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GATGCATTATTCTAATAGTTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTA 720

Qy    883 GGGACTTTTACAAACGCTAANGTTCAACCTTCAAACTACAGTTTGTTGTTCTTTTAGTGCT 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGACTTTTACAAACGCTAANGTTCAACCTTCAAACTACAGTTTGTTGTTCTTTTAGTGCT 780

Qy    943 GATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCT 1002
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    781 GATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTGAAATGGTGATGAACCT 840

Qy   1003 TCAAGTAGAATTC 1015
          |||||||||||||
Db    841 TCAAGTAGAATTC 853


RESULT 12
E02342

LOCUS       E02342              708 bp    DNA         linear      PAT 29-SEP-1997
DEFINITION  DNA sequence coding for MG-1.
ACCESSION   E02342
VERSION     E02342.1  GI:2170577
KEYWORDS    JP 1990111795-A/1.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   Kodama,K., Saito,S., Yanagida,N., Kamogawa,K., Iritani,K. and
            Aoyama,S.
  TITLE     FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE
            THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME
  JOURNAL   Patent: JP 1990111795-A 1 24-APR-1990;
            NIPPON ZEON CO LTD, SHIONOGI & CO LTD
  COMMENT   OS   Mycoplasma gallisepticum
            PN   JP 1990111795-A/1
            PD   24-APR-1990
            PF   02-JUN-1989 JP 1989140283
            PR   02-JUN-1989 JP 88P   136343
            PR   02-JUN-1988 JP 88P   136343
            PI   KODAMA KAZUMI,  SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,
            PI   IRITANI KOICHI, AOYAMA SHIGEMI
            PC   C07K13/00,C07K7/06,C07K15/04,C12N1/21,C12N15/31//A61K37/02, PC
                 C12P21/02,
            PC
G01N33/569,(C12N1/21,C12R1:19),(C12N15/31,C12R1:35),(C12P21/02, PC
C12R1:19];
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   *source: strain=S6;
            FH   Key                Location/Qualifiers
            FH
            FT   CDS                1..708
            FT                      /product='MG-1'.

FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2096"
ORIGIN

Query Match             52.1%;  Score 679.8;  DB 6;  Length 708;
Best Local Similarity   97.6%;  Pred. No. 1.9e-89;
Matches   690;  Conservative   0;  Mismatches  17;  Indels   0;  Gaps   0;

Qy    280 TGTATGTCTATTACTAAAAAAGATGCAAACCCAATAATGCCAAACCCAATTAGAAGCA 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      1 TGTATGTCTATTACTAAAAAAGATGCAAACCCAATAATGCCAAACCCAATTACCAAGCA  60

Qy    340 GCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAA 399
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     61 GCGCGAATGGAGTTAACTGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAA 120
```

```
Qy   400  GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC  459
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GACTATGCTAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC  180

Qy   460  AATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCC  519
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   181  AATAACCTTAATGCAACACTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCC  240

Qy   520  ATCAACCAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTAGTTGAA    579
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   241  ATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACATCCAAATTAGTTGAA   300

Qy   580  GCATACAAAGCACTAAAAAACCACTTTAGAACAACAACGTGCTACTAACCTTGAAGGTTTGTCA  639
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   301  GCATACAAAGCACTAAAAAACCACTTTAGAACAACAACGTGCTACTAACCTTGAAGGTTTAGCT  360

Qy   640  TCAACTGCTTATAAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGT  699
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  TCAACTGCTTATAATCAGATTCGTAATAATTTAGTGGATCTATACAATAATGCTAGTAGT  420

Qy   700  TTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTTAGATTCTAATGAGATT  759
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TTAATAACTAAAACACTAGATCCACTAAATGGGGAATGCTTTTTAGATTCTAATGAGATT  480

Qy   760  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT  819
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAAT  540

Qy   820  GCTGATGCATTATCTAATAGTTTTATTTAAAAAGTGATTCAAAATAATGAACAAAGTTTT  879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GCTGATGCATTATCTAATAGTTTTATTTAAAAAGTGATTCAAAATAATGAACAAAGTTTT  600

Qy   880  GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGT  939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTAGT  660

Qy   940  GCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTG  986
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   661  GCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGAACGGTTTG  707
```

Feature	Location/Qualifiers
rep_origin	/db_xref="taxon:233150" 1..1889 /note="putative oriC"
protein_bind	complement(24..32) /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach; Mycoplasma pulmonis consensus ttatcMaMa" /bound_moiety="DnaA" /function="initiation of the chromosome replication"
protein_bind	complement(36..44) /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach; Mycoplasma pulmonis consensus ttatcMaMa" /bound_moiety="DnaA" /function="initiation of the chromosome replication"
protein_bind	114..122 /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach" /bound_moiety="DnaA" /function="initiation of the chromosome replication"
repeat_region	209..217 /note="oriC imperfect repeat"
protein_bind	complement(254..262) /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach; Mycoplasma mycoides consensus YtWtMHaVR" /bound_moiety="DnaA" /function="initiation of the chromosome replication"
protein_bind	298..306 /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach; Mycoplasma mycoides consensus YtWtMHaMR" /bound_moiety="DnaA"
repeat_region	complement(401..450) /note="oriC imperfect repeat"
repeat_region	complement(462..510) /note="oriC imperfect repeat"
repeat_region	complement(575..623) /note="oriC imperfect repeat"
repeat_region	complement(713..761) /note="oriC imperfect repeat"
repeat_region	complement(763..811) /note="oriC imperfect repeat"
repeat_region	complement(821..869) /note="oriC imperfect repeat"
repeat_region	872..920 /note="oriC imperfect repeat"
repeat_region	991..1004 /note="oriC imperfect repeat"
repeat_region	1061..1109 /note="oriC imperfect repeat"
repeat_region	1147..1195 /note="oriC imperfect repeat"
protein_bind	complement(1185..1192) /note="putative DnaA-box; Mycoplasma spp. putative dnaA binding site; consensus sequence approach; Mycoplasma capricolum consensus ttWtMHaa" /bound_moiety="DnaA" /function="initiation of the chromosome replication"
repeat_region	complement(1230..1278) /note="oriC imperfect repeat"
repeat_region	1326..1374 /note="oriC imperfect repeat"
repeat_region	complement(1551..1564) /note="oriC imperfect repeat"
repeat_region	complement(1609..1657) /note="oriC imperfect repeat"
repeat_region	complement(1700..1748) /note="oriC imperfect repeat"
repeat_region	complement(1831..1879) /note="oriC imperfect repeat"
repeat_region	1890..2681 /note="oriC imperfect repeat"
gene	

```
RESULT 13
AE016967
LOCUS       AE016967              301042 bp    DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 1 of 4 of the complete
            genome.
ACCESSION   AE016967 AE015450
VERSION     AE016967 AE015450
            AE016967.1  GI:31541048
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 301042)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
   PUBMED   12949158
REFERENCE   2  (bases 1 to 301042)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzell,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eaglevile Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..301042
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
```

CDS
```
/gene="parA"
/locus_tag="MGA_0619"
/note="synonym: soj"
1890..2681
/gene="parA"
/locus_tag="MGA_0619"
/note="MGA_0619 ATPase involved in chromosome
partitioning; pfam00991 ParA family ATPase; COG1192 Soj
ATPases involved in chromosome partitioning, MGR_001"
/codon_start=1
/transl_table=4
/product="ParA/Soj"
/protein_id="AAP56351.1"
/db_xref="GI:31541049"
/translation="MMISFINNKGGVLKTTLATNICGVFSKFFPKSRSVIVDLDGQG
KNKDNQVIDKNDPFFYSESFEEARIQLIKDKKVEVKEEBKVQETTVKNKLSEAKKEB
NVSATFGOHPERLKNTLIDIFRGEKDIDDCVLNVFPSIDILPSNHELSFVDMDVARKE
YKLSVIKNLIEKLEEMYDFVFLDTPPAMSTIVSVSMEISDMIVIPFEPDQYSILGLLR
VIDTIETRERNEDLKVLVVPTKVNARTRLENDVLNIVKTKLSKKNIPLSKNIVSFTT
KSSASVGVEKLPIVLINQRSKYQEBYIEITKEIINLLKDNKNNEK"
```

gene
```
2659..3147
/locus_tag="MGA_0621"
```

CDS
```
2659..3147
/locus_tag="MGA_0621"
/note="MGA_0621 PS50318; MGR_002"
/codon_start=1
/transl_table=4
/product="unique hypothetical"
/protein_id="AAP56352.1"
/db_xref="GI:31541050"
/translation="MTRTMKNKKAKKKERRFTDLSADLDEBVEKIDPEYEDFKEIKIE
KNKDNQVIDKNDPFFYSESFEEARIQLIKDKKVEVKEEBKVQETTVKNKLSEAKKEB
AKDVYIDSSLEIASQEPLTKGMFYTNSRIIRRKVRECAKNKGLSISRLIITMILDKSIK
EB"
```

gene
```
3163..4548
/gene="dnaA"
/locus_tag="MGA_0622"
```

CDS
```
3163..4548
/gene="dnaA"
/locus_tag="MGA_0622"
/note="MGA_0622 ATPase involved in DNA replication
initiation COG0593"
/codon_start=1
/transl_table=4
/product="DnaA"
/protein_id="AAP56353.1"
/db_xref="GI:31541051"
/translation="MKTKLKRPLEEISVHFNEANSELLDAFVHSIDFVFBENDNIYIY
FESPYFNEFKNKLNHLINVENAVVFNDYLSLEWKKLIKENKRVNLLNKKBADTLKEK
LATLKKQEKYKINPLSKGIKEKNFGNYLVFBFNKEAVYLAKQIANKTTHSNWNPIII
EGKPGYGKSHLLQAIANERQKLFPEBKICVLSSDDFGSEFLKSVIAPDPTHIESFKSK
YKDYDLLMIDDVQIISNRPKINBTFFTIFNSLVDQKKTIVITLDCKIBEIQDKLTARM
ISRFQKGINVRLINQPKNEIIQIFPKQFKENNLEKYMDDHVIEBISDFDEGDIRKIEG
SVSTLVFMNQMYGSTKTNGQIILKSFIRKVTNNKLLISKDPKYVFDKIKYHFNVSEDV
LKSSKRKEIVQARHICMYVLKNVYNKNLSQIGKLLRKDHTVRHGIDKVEBELENDP
NLKSFLDLFKN"
```

gene
```
4744..5133
/locus_tag="MGA_1322d"
```

```
4744..5133
/locus_tag="MGA_1322d"
/note="Q COG1132 ABC-type multidrug/protein/lipid
transport system, ATPase component; similar to MGA_1284;
MGR_004"
```

CDS
```
4744..5133
/locus_tag="MGA_1322d"
/codon_start=1
/transl_table=4
/product="ABC transporter component"
/protein_id="AAP56354.1"
/db_xref="GI:31541052"
/translation="MNRKLDHCISALLYFGLINFGIIGRDKNALTQYSLILLGVAAVG
FTFGYLGGRSIIIASVEFAKQLRVNIPERYQSFSVKNTDKFEKASVLTRMTTDINFIH
QSIQSGRTAIRGMSVFLFSLVLMFVIS"
```

gene
```
5130..6077
/locus_tag="MGA_0625"
```

CDS
```
/note="synonym: mldB"
5130..6077
/locus_tag="MGA_0625"
/note="MGA_0625, MdlB-like; ABC-type
multidrug/protein/lipid transport system MdlB [Q] COG1132
PS00890; similar to MGA_1285, MGA_0626 and MGA_1287;
MGR_005"
/codon_start=1
/transl_table=4
/product="MdlB-like"
/protein_id="AAP56355.1"
/db_xref="GI:31541053"
/translation="MKLGIASVAIMPVIIGGILLVYRFVIGNYKKLFKQYDQLNNLAK
BSIAGARVVKSYHQQDNEIQKPNRVAGFIYKNFTKIERITALISPIVLFCIYALAIAI
AWIGTNNIVDGKLDIGSLASVFAYAFQMLINLLLSVVYVTIITAKPSKDRIIEVLTB
KIDIKDKKYAIDTVSDYEVEYKDVSKYVDTNPHHNLEKINIKIKKQTIGIGGSTGS
GKTSIVNLLTRLYECNEBGQVLLNNIQLNNYSIKALRDAIAIVPQKSILYSGTIKDNIL
MGGNYSDEEVEKAITQAQAABFINKLPNKLDSIVEQKWN"
```

gene
```
6369..8294
/locus_tag="MGA_0626"
```

CDS
```
6369..8294
/locus_tag="MGA_0626"
/note="MGA_0626, ABC-type multidrug/protein/lipid
transport system; similar to MdlB [Q] COG1132 PS00890;
```

```
Query Match            19.2%;  Score 250.6;  DB 1;  Length 301042;
Best Local Similarity  55.3%;  Pred. No. 5.3e-28;
Matches      587;  Conservative   0;  Mismatches 444;  Indels 30;  Gaps   4;

Qy    251 CATCTCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGCAAACC 310
              |||||     ||||| |||||    |||||||  || | ||| ||||||||| ||||||
Db 258430 CAAACCCTCCTAGTGGTTGGTATGGCGGAGATACTAATCCAGGAAATGACGGAGAA 258489

Qy    311 CAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATG 370
              |||      ||||  ||||||||||||  || ||   ||| ||||||| |||||||
Db 258490 TGGAGAATTCTGCTCAACAATTAGCAGCTGCTAAAATCATTAGGCGATTTATTAGCAA 258549

Qy    371 CTAAAGCCATGACATTAGCTTCACTACAAGAGCTGAAAGACTATGCCAAGATTGAACGC 430
              ||   |||  ||| |||||  |||||||||  ||  ||  |||| |  ||||| |
Db 258550 CACAAAGCAGTAACCTTGCAAAATATGCAGATTATCAAATATTCAGGAAATACTCTAACAG 258609

Qy    431 CTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTAATGCAACATTAGAACAACTAA 490
              ||||| || ||  | ||  | |||||   |||  |  |  || |  ||||||||||| |
Db 258610 CTGCATACCACAACTGCTAAATCTAAATCTACATCAGATAATCAAGAGCGTCACTTTAGAGCAAGTTA 258669

Qy    491 AAATGGCTAAAACTAATTAGAATCAGCCATCAACAAGCCAAGCTAATACGGATAAAACGACTT 550
              |||||| ||||   || ||| | |||| ||  ||| | || ||||||  | |||||| | ||
Db 258670 AAAGTGCTACATCACAACTTACAGCAGCTATTGATACTGCTCAAGTTCAAAAAACTAGTT 258729

Qy    551 TTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAAC 610
              |||| | | || |  |||||| |  |||||| ||  ||||  ||||| ||  ||||||
Db 258730 TTGATGAAAAAAATCCTGAATTAATCAAAGCAATATTATGCTTTAAAAGAAACATTAAAAA 258789

Qy    611 AACGTGCTACTAAACCTTGAAGGTTGTCATCAACTGCTTATAATCAAATTCGCAATAATT 670
              | ||||||||   || |||| ||| | |  || ||| | |||| |||  | |||| | |
Db 258790 ATGAAGAACAGTTTAAGTGGTTAACTGATTCTAATTTTGCAACAATCAAAACTAACC 258849

Qy    671 TAGTGGATCTATACAATACAATAAAGCTAGTAGTTTAATAACTACTAAAACACTAGATCCTGT---TA 730
              || ||| | |||| |||| |||||||  ||||| |  ||||  || |||| | |||| | |||
Db 258850 TTACTGCTTTATATCAATCTGGTAAAGATTTTGTTAAAGCAACATTAGATCGTGT---TA 258906

Qy    731 GGGGAACGCTTTTAGATTCTAATGAGATTCTACAGTTAATCGGAATATAATAATACGT 790
              | ||||||| |   ||| |||| || | |||||| | |||||| ||| ||||||| |
Db 258907 GTGGAAACGCTCCCCAAGTGCTAAAGCAGATAAAGATAATTCTAATAGTTTTATTAAAA 258966

Qy    791 TATCAACTATTAATGAACAAAAGACTAAGCTGATGCATTATCTAATAGTTTATTAAAA 850
              | |||| ||| |||||||| ||| ||||  ||||||||||| |||||||| ||||||
Db 258967 TTTCAAAACTTGAAACCTGAAAAACAAATGCTAATACTTTAGCTACAAGTTTTGTTAAAG 259026

Qy    851 AAGTGATTCAAAATAATGAACAAGTTTGTAGGGAGTTTGAGGACTTTACAAACGCTAATGTTCAAC 910
              || |||| |||||||| |||||||  |  |||||| || || ||||||| |||  ||| || |||
Db 259027 AGTTTTAGTTAAAATAAATTAACAGGAATTGATAC---AACTAATAATCGAGAACAAC 259083
```

```
Qy    911  CTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAATTAT------- 963
            |||||    ||  |||||||   || || ||||  |  |||   ||||||||||
Db 259084  CTGGAAATTACAGCTTGTTGGTTATAGTGTTAATGCACAAATAATAATGAGATTCCTA 259143

Qy    964  ----AAATATGCAAGAAGGACCGTTTGGAAATGGTGAATGAACCTTCAAGTAGAATTCTTG 1018
                |||| ||  ||| |||| ||||||||  ||| ||||| ||||  |||||||||||
Db 259144  ACTGGAATTTCGCACAAAGAAAAGGTTTGAACTAGTAGCGTAGAACTAGTCTAATTT 259203

Qy   1019  CAAACACGAATAGTA----------TCACAGATGTTCTTGGATTTATAGTTTAGGTG 1066
            ||  || || |||            ||  |||||  |   ||| |||||||||||||
Db 259204  CTAGCACTTCTGATAATTCTCTCACCTTAACAGAAGTATCAGAAGTTTATAGTTTATCAG 259263

Qy   1067  GAACAAACACAGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGGTATTTAT 1126
            |||||| |  |||||||  |||| |||  || |||| |||  |||||||||||  ||||
Db 259264  GAGCCAGGTACTAAATACAGCTTAACATTTAATTAATTACGGTCCGTCAACTGGTATTTAT 259323

Qy   1127  ATTTCCCTTATAAGTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAAT 1186
            |||  |||||||||||||| |||  ||| |||  ||| ||||||  ||  ||||| |
Db 259324  ATTTCCCTTATAAGTGGTTAAATAAACAAGGTGATGATGAAAAATAATGTAGCTTTACAATATACCT 259383

Qy   1187  TAAATAATGGAAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATACTA 1246
            |||| |  ||  ||||| ||||| || ||| ||||  ||||||  ||| |||||||| |
Db 259384  TAAATGGTGGTAGTGCACAAGAAGTTAAGTTTGCTCCAACTGTAAAAACCAGTGTTAGTG 259443

Qy   1247  CAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCT 1287
            |||| ||||||||||||||  |||  |||  || ||||||| 
Db 259444  CAGATTCTTCTGGTGATTCAAATAATCAGACAGAAGTGCT 259484
```

RESULT 14
AF210770
LOCUS   AF210770   4568 bp   DNA   linear   BCT 02-FEB-2000
DEFINITION Mycoplasma gallisepticum pMGA-like protein 9.1 gene, partial cds;
            pMGA-like protein 9.2 gene, complete cds; and pMGA-like protein 9.3
            gene, partial cds.
ACCESSION   AF210770
VERSION   AF210770.1 GI:6851355
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum
  ORGANISM Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmatales; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 4568)
  AUTHORS   Pharr,G.T., Branton,S.L., Hanson,L.A., Minion,F.C., Lott,B.D.,
            May,J.D. and Hughlett,M.B.
  TITLE   A novel pMGA-like gene from the F-strain (vaccine strain) of
            Mycoplasma gallisepticum
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 4568)
  AUTHORS   Pharr,G.T., Branton,S.L., Hanson,L.A., Minion,F.C., Lott,B.D.,
            May,J.D. and Hughlett,M.B.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-DEC-1999) College of Veterinary Medicine, Mississippi
            State University, Box 9825, MS 39762, USA
FEATURES   Location/Qualifiers
     source   1..4568
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="F"
                     /db_xref="taxon:2096"
     CDS   <1..637
                     /note="adhesin"
                     /codon_start=2
                     /transl_table=4
                     /product="pMGA-like protein 9.1"
                     /protein_id="AAF29523.1"
                     /db_xref="GI:6851356"
                     /translation="VYNDIFSNILDKQDEPNSVTVNLLKGYSLAADHSTYFYQFSSSN
                     RMNESNPTYLVGFIGGHGNRNNLNSSNTTNNEVASPSVQTSNRTLITIYVNAPKDGQYY
                     IKGSYLTSNNRNLKFTTTATANNSITFTVKGKNNWSTLGTFNTANNNDIETSGSSSSG
                     QPNESKTIKLNLNKVVITSVMMDKNPGAPYIGNLTFTLMSPTMEAKK"
     repeat_region   787..831
                     /note="microsatellite"
                     /rpt_type=tandem

```
                     /rpt_unit="gaa"
     CDS             967..2982
                     /note="adhesin"
                     /codon_start=1
                     /transl_table=4
                     /product="pMGA-like protein 9.2"
                     /protein_id="AAF29524.1"
                     /db_xref="GI:6851357"
                     /translation="MKRKNIKFVSLLGIGSFVMLAAASCTSAATPTPNPEPKPDMP
                     NPPSGDMNGGDTNPGNDGGMENSAQLAAAKKELSDLLATQSSNLAKYADYTNIQNTL
                     TAAYTPAKSTSDNTSVTLEQVKSATSTLQAAIDTAASSKTSFDEKNPELIKAYYALKE
                     TLKNEEVVLSGLTDSNFATIKTNLITALYQSGKDFVKATLDPVSGNAPQIADITKADKD
                     IADAVSKLETWKTNANTLATSFVKVLVKNKLTGIDTTNNREQPGNYSFVGYSVNATN
                     NNEIPNWNFAQRKVWTSDNGRTSLSJSSTSDNSSTLTEVSWIYSLSGAGTKYSLTFNYY
                     GPSTGYLYPFYXLVKEGDENNVALQYTLNSGSAQEVNFAPTVKTSVSADSSGDSNNQT
                     ESAAETMPVTSDLNPATVSDINIAKLTLSNLKFGSNTIEFSVPTEPSNKVAPMIGNM
                     YLTSNIANEAKVYNSIFGNVDNSSEASTYVTVDILKGYSLATNWSTYVTRPMLTNSM
                     PENATTYLVGFIGGQLARTTVGSIFNRNNFPIMNNENRTFTLYVNAPKAGDYHISGSY
                     LTRNTRGLKLIVTDTTDKNNSITITTSGKNWNTLGHPDTSKANNSNGNDGSVENNKA
                     SLTLKEGLNKLVIAGGTQDGKNAPYIGNLIFTLNNSSTNASQDSSST"
     repeat_region   3147..3182
                     /note="microsatellite"
                     /rpt_type=tandem
                     /rpt_unit="gaa"
     CDS             3313..>4568
                     /note="adhesin"
                     /codon_start=1
                     /transl_table=4
                     /product="pMGA-like protein 9.3"
                     /protein_id="AAF29525.1"
                     /db_xref="GI:6851358"
                     /translation="MKRKNIKFVSLLGIGSFVMVAAASCTSAATPTPNPEPNPNEP
                     KPDPMNPPSGGMNGGDINPGGGQNMMDSAAQELTAARTALTSLLASKNANIEMYSDY
                     AKIQNTLIAAYTAEQTSQNSSATLEQVKNATSALQTAINTANSNKQKFPQDHSNLLM
                     SYKNLMATLAKKETTWMTLKDPKYSAILDQINGVSCKGBELVQHTLDPVSGIVPAANT
                     ITEBITKIEBYISEKTLQDQKNNADQFDNYQSPTLDKTKLENVEDAKKMGQPANYSFV
                     GYSVDVTGTSGQETTIPNWNFAQRAIFTSGNQPTKVTATTGEDQSTAKPLSDVSWIY
                     SLAGTGAKYLBFTYYGPSTGWLYFPYKLVXANDDVGLQYKLNSNETILTPIFGEGTT
                     TNGPAATVENINVAKVRLTGLAFGKNT"
```

```
ORIGIN

      Query Match       19.1%;  Score 249;  DB 1;  Length 4568;
      Best Local Similarity   55.2%;  Pred. No. 3.5e-27;
      Matches  586;  Conservative   0;  Mismatches 445;  Indels  30;  Gaps   4;

Qy    251  CATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTATCAAAAAGATGCAAACC 310
            ||||||| ||||||||||| |||| |||  |||  || ||| |||| |  ||||||||||
Db   1097  CAAACCCTCCTAGTGGTGATATGAATGGCGGAGATACTAATCCAGGAAATGACGGAGGAA 1156

Qy    311  CAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGATCATCAATCAATG 370
            ||||||||||  |||  |||| || |||| || |  ||||||  || ||||| ||
Db   1157  TGGGAGAATTCGTCTCAACAATTAGCAGCTGCTAAAAAGAATTAAGCGATTTATTAGCAA 1216

Qy    371  CTAAAGCGATGACATTAGCTTCACTACAGACATGCCAAGATTGAAGCTAGTTTATCAT 430
            |||||||| || || || |||| |||| ||  ||||| |||| |||| || |||||
Db   1217  CACAAAGCAGTAACCTTGCAAAATATGCCAAAATATCAGAATATACAAATATTCAGAATACTCTAACAG 1276

Qy    431  CTGCTTATAGTGAAGCTGAAACAGTTAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAA 490
            |||  || |||||||| || ||| || || ||||  || |||| |||||| |||| ||||
Db   1277  CTGCATACAACACTGCTAAATCTACATCAGATAATACACAGGCGTCACTTTAGAGCAAGGTTA 1336

Qy    491  AAATGGCTAAAACTAATTTAGAATCACAGTAACTAATTACGGATAAAAACGACTT 550
            ||||||| ||| | |||| |||||  |  ||  ||||||||  ||||| |  ||||
Db   1337  AAAGTGCTACATCAACTTTACAGCAGCTATTGACAAGCCACTAAATAATGATGTTCAAAAAACTAGTT 1396

Qy    551  TTGATAATGAACACCAAATTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAAC 610
            |||| ||||  | ||||||| ||  |||| | |||  |||||  ||||||| |||
Db   1397  TTGATGAAAAAAATCCTGAATTAAATCAAATCAAAGCCATTATGATGGATAAGAACATTAAAAA 1456

Qy    611  AACGTGCTACTAACCTTGGAAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATT 670
            || |||| ||| |||  || |||| || ||  || || ||  ||||| |  |||||||
Db   1457  ATGAAGAAACAGTTTAAGTGGTTAACTGATTCTAATTTGCAACAATCAAAACTACC 1516
```

```
Qy    671  TAGTGGATCTATACAATAAAGCTAGTAGTTTTATTAATAACTAAAAACACTAGATCCACTAAATG  730
                ||||   |||||  |||||||                 ||||  ||||||| |||
Db   1517  TTACTGCTTTATATCAATCTGGTAAAGATTTTGTTAAAGCAACATTAGATCCTCGT---TA  1573

Qy    731  GGGGAAACGCTTTTAGATTCTAAATGAGATTACTACAGTTAATCGGAAATATTAATAATACGT  790
           || || ||||||  ||||   ||||  || || ||||| |||| |  ||||| | || |||
Db   1574  GTGGAAACGCTCCCCAAATGCTGATATTACAAAAGCAGATAAAGATATTGCTGATGCTG  1633

Qy    791  TATCAACTATTAATGAACAAAAGACTAATGCATTATCTAATAGTTTTATTAAAA  850
           ||| ||| ||||| ||| ||  || ||||||| || |  ||||| ||| ||||
Db   1634  TTTCAAAACTTGAAACCTGAAACAAATGCTAATACTTAGCTACAGAGTTTTGTTAAAG  1693

Qy    851  AAGTGATTCAAAATAATGAACAACAAGTTTGTGAGGGACTTTTACAAAGGCTAATGTTCAAC  910
           | ||| ||||||||   |||| |||   |||  ||||         |||||  |||||| |
Db   1694  AGGTTTTAGTTAAAATAAATTAACAGGAATTGATAC---AACTAATAATCGAGAACAAC  1750

Qy    911  CTTCAAACTACAGTTTGTTGCTTTAGTGCTGATGTAACACCCGTCAATTAT-------  963
           | ||||| || |||  ||  |||||| |||  |  |||| ||||||| |||
Db   1751  CTGGAAAATACAGCTTGTTGTTGGTTATAGTGTTAATGCAACAAATAATGAGATTCCTA  1810

Qy    964  -----AAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTG  1018
                |||||| |  ||  | |||||| || ||  ||||| ||||   |||||  |||||
Db   1811  ACTGGAATTTCGCACAAAGAAAGGTTGAACTAGTGATAACGGTAGAACTAGTCTAATTT  1870

Qy   1019  CAAACACGAATAGTA---------TCACAGATGTTTCTTGGATTTTATAGTTTAGCTG  1066
           |||| |||||| |||          | || |||  |  |||| |||||||||| ||||
Db   1871  CTAGCACTTCTGATAATTCTTCACCTTAACGAAGTATCATGAATTATAGTTTATTCAG  1930

Qy   1067  GAACAAACCGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTAT  1126
           ||| ||| |||  || ||  ||||||||| |  ||| ||  || |||  ||| |||| |
Db   1931  GAGCAGGTACTAAATACAGCTTAACATTTAATTAATTTATTACGGTCCTTCCACTGGTTATTTAT  1990

Qy   1127  ATTTCCCTATAAGTTGGTTAAAGCAGCAGTCGATGCTAATAACGTTGGATTAACATACAAAT  1186
           ||||||| ||||||||||||| |||||||| |||| | || ||||| ||||  ||||| ||
Db   1991  ATTTCCCATATAAGTTAGTTTAAGAAGGTGATGAAAATAATGAAATAATGAGCTTTACAATAACCT  2050

Qy   1187  TAAATAATGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATCACTA  1246
           || |||| |||| |  ||||  |||  | ||| ||| ||||  ||| ||||||  ||||
Db   2051  TAAATAGTGGTAGTGCACAAGAAGTTAATTTTGCTCCAACTGTAAAAACTAGTGTTAGTG  2110

Qy   1247  CAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCT  1287
           |||  ||||| ||| ||||  |||| ||||| ||||||| ||
Db   2111  CAGATTCTTCTGGTGATTCAAATAATCAGACGACGAAAGTGCT  2151
```

RESULT 15
AE016969/c

```
LOCUS       AE016969          301903 bp   DNA     linear   BCT 03-SEP-2003
DEFINITION  Mycoplasma gallisepticum strain R section 3 of 4 of the complete
            genome.
ACCESSION   AE016969 AE015450
VERSION     AE016969.1  GI:31541483
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum R
  ORGANISM  Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 301903)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE     The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
   PUBMED   12949158
REFERENCE   2  (bases 1 to 301903)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzell,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT
            06269-3089, USA
FEATURES             Location/Qualifiers
     source          1..301903
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:233150"
     gene            complement(77..754)
                     /locus_tag="MGA_0090"
     CDS             complement(77..754)
                     /locus_tag="MGA_0090"
                     /note="MGA_0090; [R] COG1881 Phospholipid-binding protein;
                     MGR_456"
                     /codon_start=1
                     /transl_table=4
                     /product="predicted phospholipid-binding protein"
                     /protein_id="AAP56784.1"
                     /db_xref="GI:31541484"
                     /translation="MRIMKKYSIGNNSKIWSKMVGADGYLALHGGGHKNENQTPPSF
                     SLDLSWEKIPNAKSYAVIVEDFDAAAVIGFSFIHWIAVNIKTNHIAENQSYLDHNKWL
                     ESKQSAYGENLLWQGHNSSVNKTLVANNKTNKDLGGILPEGFTSTSLENSLMYFGCYP
                     PNADHTYTVVYGLDVDASLTYHKNHRTQKYHLNKPYYVGDFMQAIHKHVGKHTLH
                     FKYKKVG"
     gene            complement(767..1396)
                     /locus_tag="MGA_0091"
     CDS             complement(767..1396)
                     /locus_tag="MGA_0091"
                     /note="MGA_0091; [R] COG1881 Phospholipid-binding protein;
                     MGR_457"
                     /codon_start=1
                     /transl_table=4
                     /product="predicted phospholipid-binding protein"
                     /protein_id="AAP56785.1"
                     /db_xref="GI:31541485"
                     /translation="MKKYISYIFNLNYYFDNYLRWIINMDMKKKHAIWSNCFTETGDG
                     VYLKSETVKLNNNRYESLDLTWDKIEGAKSYAVMVDYEASRVIGQSFIHWVANVKE
                     NKLAYAANINNKSIVQGLNSTAQQGTESSKGVIIECVPGAFKNTAEAASDYLPPLPPD
                     DTHLYTLRVYGLDVDHVDLPKHFSLTDLNQKIIRHTVGEHBLHFWYKPN"
     gene            complement(1437..2306)
                     /gene="ara1"
                     /locus_tag="MGA_0093"
     CDS             complement(1437..2306)
                     /gene="ara1"
                     /locus_tag="MGA_0093"
                     /note="MGA_0093; ARA1 [R] COG0656 Aldo/keto reductases
                     related to diketogulonate reductase; MGR_458"
                     /codon_start=1
                     /transl_table=4
                     /product="ARA1"
                     /protein_id="AAP56786.1"
                     /db_xref="GI:31541486"
                     /translation="MYLTTKKIKVMKYNLLANNYKMPSIGFGYTYKLEDGKQTVEAVKT
                     ALEVGYRHLDCAEIYENQLSVSQGIKLSGVNRKEIFITSKIWNDDKGYESTKAFHKI
                     LKDLDIBYLDLLLIHWPIGKGFKDNWQEVNAQTWKAMEEFYLEGKIKAIGLSNFIVHH
                     IEALKKSAKILPMVNQLEFHPGYLQPEIVEYCQKNNIVVQAWSPPSQSKVLENKTLVQ
                     LAAKYEVTVAQLIANWIIQKDIVPLPKSKTPBRIRSNFDVFNFGINQADLKVIDQLND
                     CGGLGLHPDQVDF"
     gene            2541..2930
                     /locus_tag="MGA_0094"
     CDS             2541..2930
                     /locus_tag="MGA_0094"
                     /note="MGA_0094; MGR_459"
                     /codon_start=1
                     /transl_table=4
                     /product="conserved hypothetical"
                     /protein_id="AAP56787.1"
                     /db_xref="GI:31541487"
                     /translation="MINLLIPTMVFNKSKDLLISKPRIYINLKSNMISKLVRVDLNSV
                     IINLQKNELNLPLNNRRPDQLLARKKK"
     gene            complement(2997..3440)
                     /locus_tag="MGA_0096"
     CDS             complement(2997..3440)
                     /locus_tag="MGA_0096"
                     /note="MGA_0096; MGR_460"
```

```
                                    LKEIKTEFSFIKQITPLEFSKQYVNYQEYFDLDLWTIGNTIDYKINQRYSLKDSNNLF
                                    EASDFNLVDRNILKPFALANLVAKQTMWFACDVNHYRNNKTGPDNQQPDYQSLFN
                                    IDFSVDRNKQVRSHFISSNHAMTLSGVDFDEAKSLLKQKELVKKNLKKPDQYQFVL
                                    DLSQTFVFKKWKIENSWGEKVGNKGFYYMNDQWFNDYLIDIVISTKAADFFKNPKPI
                                    DPDLIGLVKQLKTKEILKAGLDTKYPILIDGYDPLGANNKGVK"
        gene        complement(6968..7402)
                    /gene="smpB"
                    /locus_tag="MGA_0102"
        CDS         complement(6968..7402)
                    /gene="smpB"
                    /locus_tag="MGA_0102"
                    /note="MGA_0102; SmpB [O] COG0691 tmRNA-binding protein;
                    MGR_465"
                    /codon_start=1
                    /transl_table=4
                    /product="SmpB"
                    /protein_id="AAP56793.1"
                    /db_xref="GI:31541493"
                    /translation="MMRLLVNNKKAKYNYELLDKYEAGISLSGNEVKSLALHHGKLDD
                    SYVIIRKNEAYLLNLLIPKYKFDTSKVLNETRTRKLLLHKSEILKIDLIKKQHSLVII
                    PYQIYFVNNKIKVSYILARPKKRYDKRQTKEREINKKIRKY"
        gene        complement(7399..8148)
                    /locus_tag="MGA_0103"
        CDS         complement(7399..8148)
                    /locus_tag="MGA_0103"
                    /note="MGA_0103; MGR_466"
                    /codon_start=1
                    /transl_table=4

Query Match          17.1%;     Score 222.8;     DB 1;     Length 301903;
Best Local Similarity 56.2%;     Pred. No. 5.7e-24;
Matches   516;     Conservative    0;     Mismatches 372;     Indels   30;     Gaps    4;

Qy     323  AAACCCAATTAGAAGCAGCGCGAAATGAGAGTTAACAGATCTAATCAATCAATGCTAAAGCGATGA   382
            || || |||||||| | |||| ||| || ||||| ||| || | ||||||||||||| ||||| ||
Db  220537  AACAACAATTAGCAACTGCTAGAAAAACGCTAACCGATCTAATTGGTACACGAAAATACCA    220478

Qy     383  CATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTG   442
            |||||| ||| ||| ||| ||||| |||||| ||||||  |||  ||||| ||||| |||
Db  220477  ACGTTGCTTATATGCTGATTATGCCAAAATTCAAAGCACTTTAAGCACTGCTTATATGA    220418

Qy     443  AAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAA   502
            |||||||| |||| || |||| ||||||| ||||||||||||  | || |||| ||||||
Db  220417  CAGCTAAAACTGCATCAGCAGAAAATACAAATGCAACTTTAGAAAAATCTAAGATCTGAAC    220358

Qy     503  CTAATTAGAAATCAGCCATCAACAAGCTAATACGGATAAAACGACTTTGATAATGAAC   562
            |||||||||||| ||| |||| ||| |||||||||| || ||| |||  |||||||| ||
Db  220357  CTACACTACAAACAGCAATTGATAAAGCTGTTAGTAGTAAAACCACTTAGAACAACGTGCTACTA    220298

Qy     563  ACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTA   622
            ||||||||| ||||||| | || ||||| | ||||||  || ||||||| ||||| |||
Db  220297  ATGCTGATTTAGTTAGTGCATATACATTATCAATTAAAAGATGCTGTTAAATCAGAAACTACTA    220238

Qy     623  ACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAAATTCGCAATAATTAGTGGATCTAT   682
            ||||| || | | || |||||| ||  ||||| |||||| |||||| ||||||| |||||
Db  220237  ATTTAGAAGGTTTTATCTGATGCAAATTATACAGCCATAAAAGATAACCTTACTACTTATT    220178

Qy     683  ACAATAAAGCTAGTAGTTTAATAACTAAAACAACTAAAATGGGGAACGCTTT   742
            || | |||||| |||| ||||| ||| |||| ||| ||||| |||| ||| ||
Db  220177  ATGGTAAAGCTAAGAATATTTTAACAGCTACGTACGTATATCCTGAATCAGGAGGGAGATTCCTA    220118

Qy     743  TAGATTCTAATAGATTACTACAGTTAATCACAGTATAAACAATTTTATTACAACATATTA   802
            ||||  ||| |  |  ||||  ||| || |||||||| |||||||||| ||| |||||||
Db  220117  AAG--TTGATGCTGTAAAAGTAGTAAACAATTTATTACAAGTTCAAAATGCAAATAGTAATATTG    220061

Qy     803  ATGAACAAAGACTAATGCTGATGCATTATCCAATATCTAATAGTTTATTAAAAGTGATTCAAA   862
            |||||| || |||||||| |||| |||  |||||| ||| ||||||||| |||| ||| ||||
Db  220060  AACAATGAAAAAATAATGCAAACATGTATCTGATAGTTATCTGATAGTTTTAAAGAAAAACCTTAGATA    220001

Qy     863  ATAATGAACAAAGTTTGTAGGGACTTTACAAACGACTTTCAACCTTCAACCTTCAAACTACA   922
            | ||||| |||||   ||  || | || |||  |  |||  ||||||||        ||||||
Db  220000  AAGAACAATTAACAAGTAATGCGGAATTACAA----------------CAACCTGCAAATTACA    219953
```

```
Qy       923 GTTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCG 982
             ||||||| || || ||| |  ||| ||||     |||||||  || ||| |  |||| |||
Db    219952 GTTTTGTAGCATATAATCAAGATATAACTAATCCTACTTATAATTTTGCTAAAAGAGTAG 219893

Qy       983 TTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGC--------AAACACGAATAGTA 1033
             |||||| ||  ||| |  ||  |||
Db    219892 TTTGAAAACCACAAGAAGGTCGGTCAAGTACCTATGTGCCATTAGAAAATCAAGGTGATT 219833

Qy      1034 TCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAACGAAGTACCAATTTAGTT 1093
             || |||| | ||  |  ||| || ||||| |||||| ||| |  |  || |  ||| ||
Db    219832 TAACAGACGTATCATGAATTTATAGTTTAGCTGGAAATGAAACTAAATATAGTTTTACTT 219773

Qy      1094 TTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAG 1153
             |||  | || | ||| ||  | | | |  ||||||| |||| | ||||| ||| |||
Db    219772 TTGCTAATTATGGCACAACTACAGGATACTTATATTTTCCATACAAGTCAGTTAAATCAA 219713

Qy      1154 CTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAATGTTCAACAAGTTG 1213
             ||||||||||| | || ||||| || |||| |||  || |  ||| || | || |  | |
Db    219712 GCG------ATAATGTTGCATTACAAATATAAGTTAAATGGTGCTAGTCCGGTATCTATCG 219659

Qy      1214 AGTTTGCCACTTCAACTA 1231
             |||| |||  ||   |||
Db    219658 ATTTTAATAACTCAGCAA 219641
```

Search completed: May 5, 2004, 17:06:39
Job time : 4902.91 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May  5, 2004, 09:54:21 ; Search time 457.318 Seconds
                                    (without alignments)
                                    12131.925 Million cell updates/sec

Title:       US-09-901-572A-1
Perfect score: 1306
Sequence:    1 aaaaacatcagattgttaat........taaaatcgttttatcaggtt 1306

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:   3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:   6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_29Jan04:*
             1: geneseqm1980s:*
             2: geneseqm1990s:*
             3: geneseqm2000s:*
             4: geneseqm2001as:*
             5: geneseqm2001bs:*
             6: geneseqm2002s:*
             7: geneseqm2003as:*
             8: geneseqm2003bs:*
             9: geneseqm2003cs:*
            10: geneseqm2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the total score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1282	98.2	2346	2	AAQ77857	Aaq77857 Mycoplasm
2	1264	96.8	1387	2	AAQ77854	Aaq77854 Mycoplasm
3	1262.4	96.7	1387	2	AAQ53419	Aaq53419 Mycoplasm
4	1260.8	96.5	1387	2	AAQ94711	Aaq94711 Mycoplasm
5	1086.8	83.2	1935	2	AAQ77856	Aaq77856 Mycoplasm
6	1028	78.7	1152	7	ACF03364	Acf03364 Mycoplasm
7	1024.2	78.4	1371	2	AAT96595	Aat96595 Hybrid Ma
8	1022.6	78.3	3261	2	AAT96596	Aat96596 Hybrid Ma
9	989.4	75.8	1015	2	AAT04076	Aat04076 M.gallise
10	921.6	70.6	1082	7	ACF03382	Acf03382 M. gallis
11	824.2	63.1	853	1	AAQ04687	Aaq04687 DNA encod
12	801.8	61.4	853	1	AAN92574	Aan92574 DNA seque
13	679.8	52.1	708	1	AAN92568	Aan92568 DNA seque
14	678.2	51.9	708	1	AAQ04686	Aaq04686 DNA encod
15	216.6	16.6	2196	2	AAT04075	Aat04075 M.gallise
16	216.6	16.6	2369	2	AAQ77853	Aaq77853 Mycoplasm
17	205.6	15.7	2452	2	AAQ68670	Aaq68670 PMGA1.3 M
18	205.6	15.7	2453	2	AAT51536	Aat51536 Mycobacte
19	183.8	14.1	2416	2	AAQ68671	Aaq68671 PMGA1.4 M
20	183.2	14.0	2417	2	AAT51537	Aat51537 Mycobacte
21	174.8	13.4	2809	2	AAT51531	Aat51531 Mycobacte
22	173.2	13.3	2282	2	AAT51535	Aat51535 Mycobacte
23	173	13.2	2881	2	AAQ68668	Aaq68668 PMGA 1.2
24	160.6	12.3	2004	2	AAT51538	Aat51538 Mycobacte
25	150.8	11.5	2337	2	AAQ68669	Aaq68669 PMGA1.2 M
26	149.2	11.4	2577	2	AAQ68672	Aaq68672 Partial P
27	136.4	11.4	702	1	AAQ05650	Aaq05650 DNA encod
28	136.4	10.4	702	1	AAN92570	Aan92570 DNA seque
29	78.2	6.0	594	1	AAQ05651	Aaq05651 DNA encod
30	78.2	6.0	594	1	AAN92571	Aan92571 DNA seque
31	73.4	5.6	4985	6	ABQ75107	Abq75107 Anopheles
32	73.4	5.6	4985	9	ACF79720	Acf79720 Mosquito
33	73	5.6	6292	4	AAS46735	Aas46735 Tumour su
34	69.6	5.3	7892	6	ABK40056	Abk40056 Human che
35	68.6	5.3	8136	6	ABK39957	Abk39957 Human che
36	68.6	5.3	8136	6	ABL32555	Abl32555 Human imm
37	68.4	5.2	11836	4	AAS45395	Aas45395 Chemical1
38	68.4	5.2	11836	6	ABK28240	Abk28240 DNA trans
39	68.2	5.2	8392	6	ABL33490	Abl33490 Human imm
40	68	5.2	12405	4	AAS45330	Aas45330 Chemical1
41	68	5.2	12405	6	ABK28169	Abk28169 DNA trans
42	68	5.2	12405	6	AAS61143	Aas61143 Human gen
43	67	5.1	8076	6	ABK39955	Abk39955 Human che
44	66.6	5.1	8056	7	ABZ10246	Abz10246 Haematopo
45	66.4	5.1	6669	6	ABL32193	Abl32193 Human imm

ALIGNMENTS

```
RESULT 1
AAQ77857
ID   AAQ77857 standard; DNA; 2346 BP.
XX
AC   AAQ77857;
XX
DT   25-MAR-2003  (revised)
DT   23-JUN-1995  (first entry)
XX
DE   Mycoplasma gallisepticum antigen TM-67 coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   CDS            202..2049
FT                  /*tag= a
FT                  /product= "Mycoplasma_antigen"
FT   mat_peptide    202..2046
FT                  /*tag= b
FT                  /codon= seq:TGA,aa:Trp
XX
PN   WO9423019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994;   94WO-JP000541.
XX
PR   31-MAR-1993;   93JP-00074139.
PR   30-SEP-1993;   93JP-0245625.
XX
PA   (JAPG )  NIPPON ZEON KK.
PA   (SHIO )  SHIONOGI & CO LTD.
XX
PI   Saito S,  Ohkawa S,  Saeki  S,  Ohsawa I,  Funato H,  Iritani Y;
PI   Aoyama S,  Takahashi K;
XX
DR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63230.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.
XX
PS   Claim 4; Page 87-91; 123pp; Japanese.
```

XX
CC   A restriction fragment of the insert of M.gallisepticum genomic clone pUM
CC   -67 containing an open reading frame was sequenced (AAQ77857). The ORF
CC   encodes an antigenic polypeptide. A recombinant avipox virus comprising
CC   the coding sequence can be used as a live vaccine to protect against
CC   infection by Mycoplasma gallisepticum. (Updated on 25-MAR-2003 to correct
XX   PN field.)
SQ   Sequence 2346 BP; 863 A; 385 C; 360 G; 731 T; 0 U; 7 Other;

Query Match   98.2%; Score 1282; DB 2; Length 2346;
Best Local Similarity   98.9%; Pred. No. 9.8e-199;
Matches 1291; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy    1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAACACAAAATCTTCTAACAA  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AAAAACATCAGATTGTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA  60

Qy   61 AATCCTAAATAAATAAGCCGTAAATTAACTAACTAAAAAAATTAAAAAAAATGGTTTTCTTTATC 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AATCCTAAATAAATAAGCCGTAAATTAACTAACTAAAAAAT..AAAAAAAATGGTTTTCTTTATC 120

Qy  121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT 180

Qy  181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAAATCATCTTAAAGACTATTAGTTTG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ATAAATATATCTTAATATTCTATGAATAAGAAAGAAATCATCTTAAAGACTATTAGTTTG 240

Qy  241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTTAGCTGTGTATGTCTATTACTAAAAAA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTTAGCTGTGTATGTCTATTACTAAAAAA 300

Qy  301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGTAACAGAT 360
        ||||| |||||||||||||||||||||||||||||| ||||||||||||||||||| ||
Db  301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCGCGGAATGGTAACTGAT 360

Qy  361 CTTAATCAATGCTAAAGCGATGACATAGCTTCACTACACACGACTATGCCAAGATTGAAGCT 420
        |||||||||||||||||| || ||||| ||||||||||| ||||||||| ||||||||||||
Db  361 CTTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAAGACTATGCTAAGATTGAAGCT 420

Qy  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA 480

Qy  481 GAACAACTAAAAATGGCTAAAAAAATTAGAATCAGCGTCAACCAAGCTAAATAACGAT 540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAACAACTAAAAATGGCTAAAAAAATTAGAATCAGCGTCAACCAAGCTAAATAACGAT 540

Qy  541 AAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGACTAAAAACC 600
        ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  541 AAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGACTAAAAACC 600

Qy  601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAGATT 660

Qy  661 CGCAATAATTAGTGGATCTATACAATAAATGCTAGTAGTTTAATAACTAAAACACTAGAT 720
        || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 CGTAATAATTAGTGGATCTATACAATAAATGCTAGTAGTTTAATAACTAAAACACTAGAT 720

Qy  721 CCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  721 CCACTAAATGGGGAAATGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy  781 AATAATACGTTATCAACTATTAATGACAAAAGACTAATGCTGATGCATTATCTAATAGT 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATAATACGTTATCAACTATTAATGACAAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy  841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGGACTTTTACAAACGCT 900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGGACTTTTACAAACGCT 900

Qy  901 AATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTTAGTGCTGATGGTAACACCGTCAAT 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTTTAGTGCTGATGGTAACACCGTCAAT 960

Qy  961 TATAAATATGCAAGAAGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TATAAATATGCAAGAAGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy 1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTTAGCTGGAACAAACACGAAG 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 AACACGAATAGTATCACAGATGTTCTTGGATTATAGTTTAGCTGGAACAAACACGAAG 1080

Qy 1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140

Qy 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260

Qy 1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
        ||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
```

RESULT 2
AAQ77854
ID   AAQ77854 standard; DNA; 1387 BP.
XX
AC   AAQ77854;
XX
DT   25-MAR-2003 (revised)
DT   23-JUN-1995 (first entry)
XX
DE   Mycoplasma gallisepticum 40kD antigen coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma 40kD antigen; TTM-1;
KW   ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key   Location/Qualifiers
FT   CDS   202..1308
FT   /tag= a
FT   /product= "40kD antigen"
FT   /note= "ORF includes 2 NNN codons"
XX
PN   WO9443019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994; 94WO-JP000541.
XX
PR   31-MAR-1993; 93JP-00074139.
PR   30-SEP-1993; 93JP-00245625.
XX
PA   (JAPG) NIPPON ZEON KK.
PA   (SHIO) SHIONOGI & CO LTD.
XX
PI   Saito S, Ohkawa S, Saeki S, Ohsawa I, Funato H, Iritani Y;
PI   Aoyama S, Takahashi K;
XX
PR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63227.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.

```
XX
PS   Claim 4; Page 71-74; 123pp; Japanese.
XX
CC   The plasmid pUTTM-1P contains a sequence (the TTM-1 gene) coding for the
CC   40kD antigen of Mycoplasma gallisepticum under the control of a synthetic
CC   promoter. A 1300 bp restriction fragment containing the promoter-ORF
CC   sequence was excised and was used in the construction of plasmid pNZ7929-
CC   R2. This in turn was involved in the construction of a recombinant avipox
CC   virus vector comprising the TTM-1 gene, DNA encoding the signal membrane
CC   anchor peptide from Newcastle Disease Virus haemagglutinin neuraminidase
CC   and FPV sequences. The recombinant avipox virus is useful as a live
CC   vaccine to protect against infection by Mycoplasma gallisepticum.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1387 BP; 531 A; 220 C; 203 G; 427 T; 0 U; 6 Other;

Query Match            96.8%;   Score 1264; DB 2; Length 1387;
Best Local Similarity  98.3%;   Pred. No. 8.2e-196; Indels 1; Gaps 1;
Matches 1284; Conservative   0; Mismatches 21;
```

RESULT 3
```
AAQ53419
ID   AAQ53419 standard; DNA; 1387 BP.
XX
AC   AAQ53419;
XX
DT   25-MAR-2003 (revised)
DT   16-JUN-1994 (first entry)
XX
DE   Mycoplasma gallisepticum antigen coding sequence.
XX
KW   Vaccine; mycoplasma infection; poultry; fowl; ds.
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   misc_difference 51
FT                   /*tag= b
FT                   /note= "only partially legible in specification; either C
FT                   or G"
FT   misc_difference 54
FT                   /*tag= c
FT                   /note= "only partially legible in specification; either C
FT                   or G"
FT   misc_difference 58
FT                   /*tag= d
FT                   /note= "only partially legible in specification; either C
FT                   or G"
FT   misc_difference 64
FT                   /*tag= e
FT                   /note= "only partially legible in specification; either C
FT                   or G"
FT   CDS             202..1308
FT                   /*tag= a
FT                   /product= "40kD antigen"
FT                   /note= "ORF includes 2 NNN codons"
XX
```

```
PN   WO9324646-A1.
XX
PD   09-DEC-1993.
XX
PF   28-MAY-1993;   93WO-JP000715.
XX
PR   29-MAY-1992;   92JP-00138819.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S, Ohkawa S, Fujisawa A, Iritani Y, Aoyama S;
XX
DR   WPI; 1993-405837/50.
DR   P-PSDB; AAR44493.
XX
PT   Mycoplasma gallisepticum antigen and DNA coding for it - useful for
PT   vaccination of fowl against mycoplasma infections.
XX
PS   Claim 2; Page 23-26; 37pp; Japanese.
XX
CC   The sequence coding for the 40kDa antigen was obtained by PCR
CC   amplification of M.gallisepticum genomic DNA. The antigen reacts with
CC   Mycoplasma-immune or Mycoplasma-infected serum and can be used as a
CC   vaccine to protect fowl from M.gallisepticum infection. (Updated on 25-
CC   MAR-2003 to correct PN field.)
XX
SQ   Sequence 1387 BP; 531 A; 216 C; 203 G; 427 T; 0 U; 10 Other;

Query Match      96.7%;   Score 1262.4;   DB 2;   Length 1387;
Best Local Similarity   98.0%;   Pred. No. 1.5e-195;
Matches 1280;    Conservative   4;   Mismatches 21;   Indels   1;   Gaps   1;

Qy     1  AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAAATCTTCTAACAA   60
          ||||||||||||||||||||||| || || |||||||||| || ||| || : |||: |
Db     1  AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAAATSTSTAASAA   60

Qy    61  AATCCTAAATAAATAAGCCGTTAAATAACTAAAAAAATTAAAAAAAATGGTTTTTCTTATC  120
          |||: |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AATSCTAAATAAATAAGCCGTTAAATAACTAAAAAAATTAAAAAAAATGGTTTTTCTTATC  120

Qy   121  AACCAAAAATTCTCTAGTAATAAACGCTTATTTTATTTTTTTAGTCATCTTTTAAGAT  180
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   121  AACCAAAAATTCTCTAGTAATAAACGCTTAAATTAATTTTTATTTTTTTAGTCATCTTTTAAGAT  180

Qy   181  ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240

Qy   241  TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTGTATTACTACTAAAAAA  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTGTATTACTACTAAAAAA  300

Qy   301  GATGCAAACCCAAATAATGGCCAAACCCAATTAGAGCAGCAGCGCGAATGGAGTTAACAGAT  360
          || ||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||
Db   301  GACGCAAACCCAAATAATGGCCAAACCCAATTACAGCAGCAGCGCGAATGGAGTTAACTGAT  360

Qy   361  CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT  420
          ||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db   361  CTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT  420

Qy   421  AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACATAACCTTAATGCAACATTA  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   421  AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACATAACCTTAATGCAACACTA  480

Qy   481  GAACAACTAAAAAATGCTAAAAACTAATTTGAATCAGCCATCAACCAAGCTAATACGGAT  540
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   481  GAACAACTAAAAAATGCTAAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT  540

Qy   541  AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC  600
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||
Db   541  AAAACGACTTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAAGCACTAAAAAACC  600

Qy   601  ACTTTAGAACAACGTGCTACTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT  660
          |||||||| ||||||||||||||||||||||||||||| |||||||||||||||||| |||
Db   601  ACTTTAGACACAACGTGCTACTACTAACCTTGAAGGTTTCAACTGCTTATAATCAGATT  660

Qy   661  CGCAATAATTAGTGATCATACATAAACTAGTAGTTAATAACTAAAACACTAGAT  720
          || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  CGTAATAATTAGTGATCATACATAAACTAGTAGTTAATAACTAAAACACTAGAT  720

Qy   721  CCACTAAATGGGGACGCTTCTAATGAGATTAGTACTACAGTTAATCGGAATATT  780
          |||||||||||||| |||| |||||||||| |||||||||||||||||||||||
Db   721  CCACTAAATGGGGAATGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT  780

Qy   781  AATAAACGTTATCAACTAATAATGAACAAAAGACAAATGCTGATGCATTATCTAATAGT  840
          ||||| ||||||||||||| |||||| ||| |||||||||||||||||||||||||||
Db   781  AATAATACGTTATCAACTATTAATGACAAAAAGACAAATGCTGATGCATTATCTAATAGT  840

Qy   841  TTTATTAAAAAAGTGATTCAAACTACAGTTTGTCTTTAGTGCTGTAACCCGTCAAT  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  TTTATTAAAAAAGTGATTCAAACTACAGTTTGTCTTTAGTGCTGTAACACCCGTCAAT  900

Qy   901  AATGTTCAACCTTCAAACTACAGTTTGTGCTTTGATGATGTAACCCGTCAAT  960
          |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   901  AATGTTCAACCTTCAAACTACAGTTTGTGCTTTGGAATGGTGATGTAAC  960

Qy   961  TATAAATATGCAAGAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020
          ||||||||||||||||||||||||          |||||||||||||||||||||||||
Db   961  TATAAATATGCAAGAGGACCGTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA  1020

Qy  1021  AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAACACGAAG  1080
          ||||||||||||||||||||||||||| ||          |||||||||||||||||||
Db  1021  AACACGAATAGTATCACAGATGTTTCTNNNNATTTATAGTTTAGCTGGAACAAACACGAAG  1080

Qy  1081  TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATTTCCTTATAAG  1140
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  1081  TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATTTCCTTATAAAG  1140

Qy  1141  TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT  1200
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||| ||
Db  1141  TTGGTTAAAGCAGCTGATGCTAATAACGTTGATTACAATACAAATTAAATAATGGAAAT  1200

Qy  1201  GTTCAACAGTTGAGTTGCCACTTGCCACTTGCCACTAGTGCAAATAATTACTACAGCTAATCCAACT  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GTTCAACAGTTGAGTTGCCACTTGCCACTTGCCACTAGTGCAAATAATTACTACAGCTAATCCAACT  1260

Qy  1261  CCAGCAGTTGATGAGATTAAAGTTTTATCAGGTT  1306
          |||||||||||||||||||||||||||||||||
Db  1261 -CAGCAGTTGATGAGATTAAAGTTTTATCAGGTT  1305
```

```
RESULT 4
AAQ94711
ID   AAQ94711 standard; DNA; 1387 BP.
XX
AC   AAQ94711;
XX
DT   25-MAR-2003 (revised)
DT   12-MAR-1996 (first entry)
XX
DE   Mycoplasma gallisepticum antigenic protein TTM-1 gene.
XX
KW   Antigenic protein; vaccine; poultry; diagnosis; TTM-1; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             202..1308
FT                   /*tag= a
XX
PN   JP07133295-A.
XX
PD   23-MAY-1995.
XX
PF   27-AUG-1993;   93JP-00213102.
```

```
XX
PR   27-AUG-1993;   93JP-00213102.
XX
PA   (JAPG )  NIPPON ZEON KK.
PA   (SHIO )  SHIONOGI & CO LTD.
XX
DR   WPI; 1995-220782/29.
DR   P-PSDB; AAR76955.
XX
PT   A new antigenic protein which reacts with Mycoplasma gallisepticum - is
PT   useful in a component vaccine for use against poultry infected with M.
PT   gallisepticum.
XX
PS   Claim 6; Fig 5-6; 33pp; Japanese.
XX
CC   AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein TTM-
CC   1. TTM-1 can be used as a vaccine for M. gallisepticum infectious
CC   diseases in poultry, and as a diagnostic agent for M. gallisepticum.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1387 BP; 531 A; 222 C; 427 T; 201 G; 0 U; 6 Other;

Query Match        96.5%;     Score 1260.8;   DB 2;   Length 1387;
Best Local Similarity   98.2%;     Pred. No. 2.7e-195;
Matches 1282;    Conservative    0;   Mismatches    23;   Indels    1;   Gaps    1;

Qy      1  AAAAACAATCAGATTGTTAATCTGATAATCTTGCTTAAAAAAACACAAAATCTTCTAACAA    60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AAAAACAATCAGATTGTTAATCTGATAATCTTGCTTAAAAAAACACAAAATCTTCTAACAA    60

Qy     61  AATCCTAAATAAATAAAGCCGTTAAATTAACTAACTAAAAAAAATTAAAAAAATGGTTTTTCTATC   120
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     61  AATCCTAAATAAATAAAGCCGTTAAATTAACTAACTAAAAAAAATTAAAAAAATGGTTTTTCTATC   120

Qy    121  AACCAAAATTCTCTAGTAATAACGCTTATTTATTTTTTATTTTTAGTCATCTTTTAAGAT   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AACCAAAATTCTCTAGTAATAACGCTTATTTATTTTTTATTTTTAGTCATCTTTTAAGAT   180

Qy    181  ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG   240

Qy    241  TTAGGTACAACATCCTTTCTTAGCATTGGGATTCTCAGCTGTATGTCTATTACTAAAAAA   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  TTAGGTACAACATCCTTTCTTAGCATTGGGATTCTCAGCTGTATGTCTATTACTAAAAAA   300

Qy    301  GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGGCGAATGGAGTTAACAGAT   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  GACGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGGCGAATGGAGTTAACTGAT   360

Qy    361  CTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT   420
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    361  CTAATCAATGCTAAAGCAAGACAAGACACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT   420

Qy    421  AGTTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA   480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  AGTTTATCATCTGCTTATATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA   480

Qy    481  GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAATAACGGAT   540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACAAGCTAATAACGGAT   540

Qy    541  AAAACGACTTTTGATAATGAACACCCCAAATTAGTTGAAGCATAACAAAGCACTAAAAACC   600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  AAAACGACTTTTGATAATGAACACCCCAAATTAGTTCAACCATACAAAGCACTAAAAACC   600

Qy    601  ACTTTAGACAACGTGCTACTACCTTGAGGTTTGTCATCAACTGCTTATAAATCAAATT   660
           ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db    601  ACTTTAGACAACGTGCTACTACCTTGAGGTTTCACTGCTTCAACTGCTTATAATCAGATT   660

Qy    661  CGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGTTAATAACTAAAAACACTAGAT   720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  CGTAATAATTTAGTGGATCTATACAATAATGCTAGTAGTTAATAACTAAAAACACTAGAT   720
```

```
Qy    721  CCACTAAATGGGGAACGCTTTAGATTCTAATGATAGATTACTACAGTAGTTAATCGGAATATT   780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  CCACTAAATGGGGAATGCTTTAGATTCTAATGATAGATTACTACAGTAGTTAATCGGAATATT   780

Qy    781  AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAATAGT   840

Qy    841  TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTGAGGACTTTACAAACGCT   900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TTTATTAAAAAGTGATTCAAAATAATAATGAACAAAGTTTGTGAGGACTTTACAAACGCT   900

Qy    901  AATGTTCAACCTTCAAACTACAGTTTGTTGTTGTTCTTTAGTGCTGATGTAACACCCGTCAAT   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  AATGTTCAACCTTCAAACTACAGTTTGTTGTTGTTCTTTAGTGCTGATGTAACACCCGTCAAT   960

Qy    961  TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA   1020
           ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db    961  TATAAATATGCAAGAAGGACCGTTNNNAATGGTGATGAACCTTCAAGTAGAATTCTTGCA   1020

Qy   1021  AACACGGATAGTATCACAGATGTTTCTGGATTATAGTTTAGCTGGAACAAACACGAAG   1080
           ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   1021  AACACGGATAGTATCACAGATGTTTCTNNNATTATAGTTTAGCTGGAACAAACACGAAG   1080

Qy   1081  TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG   1140

Qy   1141  TTGGTTAAAGCAGCTGATGCTAATACGTTCGATTACAATACAAATTAAATAATGGAAAT   1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  TTGGTTAAAGCAGCTGATGCTAATACGTTGGATTACAATACAAATTAAATAATGGAAAT   1200

Qy   1201  GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATATACTACAGCTAATCCAACT   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATAATACTACAGCTAATCCAACT   1260

Qy   1261  CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT   1306
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1261  -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGTT   1305
```

```
RESULT 5
AAQ77856
ID   AAQ77856 standard; DNA; 1935 BP.
XX
AC   AAQ77856;
XX
DT   25-MAR-2003  (revised)
DT   23-JUN-1995  (first entry)
XX
DE   Mycoplasma gallisepticum antigen TM-66 coding sequence.
XX
KW   recombinant avipox virus; live vaccine; mycoplasma antigen; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..1886
FT                   /tag= a
FT                   /product= "Mycoplasma_antigen"
FT   mat_peptide     54..1883
FT                   /tag= b
FT                   /codon= seq:TGA,aa:Trp
XX
PN   WO9423019-A1.
XX
PD   13-OCT-1994.
XX
PF   31-MAR-1994;  94WO-IP000541.
XX
PR   31-MAR-1993;  93JP-00074139.
PR   30-SEP-1993;  93JP-00245625.
```

```
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Saito S, Ohkawa S, Saeki S, Ohsawa I, Funato H, Iritani Y;
PI   Aoyama S, Takahashi K;
XX
DR   WPI; 1994-333181/41.
DR   P-PSDB; AAR63229.
XX
PT   Recombinant avipox virus combining DNA encoding a polypeptide -
PT   exhibiting antigenicity of mycoplasma, useful for the production of a
PT   live vaccine.
XX
PS   Claim 4; Page 78-81; 123pp; Japanese.
XX
CC   A restriction fragment of the insert of M.gallisepticum genomic clone pUM
CC   -66 containing an open reading frame was sequenced (AAQ77856). The ORF
CC   encodes an antigenic polypeptide. A recombinant avipox virus comprising
CC   the coding sequence can be used as a live vaccine to protect against
CC   infection by Mycoplasma gallisepticum. (Updated on 25-MAR-2003 to correct
CC   PN field.)
XX
SQ   Sequence 1935 BP; 728 A; 314 C; 307 G; 586 T; 0 U; 0 Other;

Query Match             83.2%;  Score 1086.8; DB 2;  Length 1935;
Best Local Similarity   96.7%;  Pred. No. 3.8e-167;
Matches 1120;  Conservative    0;  Mismatches  37;  Indels    1;  Gaps    1;

Qy     150 TTTATTTTTATTTTTAGT-CATCTTTAAGATATAAATATATCTTAAATATTCTAATGAATA 208
           |||||||| ||||||| |  |||||||||||||||||| |||||||||||||||||||||||
Db       1 TTTATTTTTATTTTTGGTAAATCTTTAAAATATAAATATAATTTTAATATTCTAATGAATA 60

Qy     209 AGAAAAGAATCATCTTAAAGACTATTAGTTGTTGTAGGTACAACATCCTTTCTTAGCATTG 268
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAAAAAGAATCATCTTAAAGACTATTAGCTGTTAGTTGTTGTAGGTACAACATCCTTTCTTAGTATTG 120

Qy     269 GGATTTCTAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC 328
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGATTTCTAGCTGTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC 180

Qy     329 AATTAGAAGCAGGCGAATGGAGTTAACAGATCCTAATCAATGCTAAAGCGATGACATTAG 388
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AATTAGAAGCAGGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAG 240

Qy     389 CTTCACTACAAGACTATGCCAAGATGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTG 448
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCACTACAAGACTATGCCAAGATGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTG 300

Qy     449 AAACAGTTAACAATAACCTTAATCAACATTAGAACAACTAAAAAATGGCTAAAACTAATT 508
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AAACAGTTAACAATAACCTTAATCAACATTAGAACAACTAAAAAATGGCTAAAACTAATT 360

Qy     509 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAA 568
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAA 420

Qy     569 ATTTAGTTGAAGCATACAAAGCACTAAAACCACTTTAGAACAACGTGCTACTAACCCTG 628
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCCTG 480

Qy     629 AAGGTTTGTCATCAACTGCTTATAATCAAATTGCAATAATTTAGTGGATCATACAATA 688
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAGGTTTGTCATCAACTGCTTATAATCAAATTGCAATAATTTAGTGGATCATACAATA 540

Qy     689 AAGCTAGTAGTTAATAACTAAAACACTAGATCCACTAAAATGGGGAACGGCTTTTAGATT 748
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AAGCTAGTAGTTAATAACTAAAACACTACTAAAATGGGGAACGGCTTTTAGATT 600

Qy     749 CTAATGAGATTACTACAGTTAATCCGAATATTAATAATACGTTATCAACTATTAATGAAC 808
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTAATGAGATTACTACAGCTAATAAGAATATTAATAATACGTTATCAACTATTAATGAAC 660
```

```
Qy     809 AAAAGACTAATGCTGATGCATTATCTAATATAGTTTTATTAAAAAGTGATTCAAAATAATG 868
           || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AAAGACTAATGCTGATGCATTAGCTAATAGTTTTATTAAAAGAAGTGATTCAAAATAATA 720

Qy     869 AACAAAGTTTGTAGGGACTTTACAAACGCTAAGTCAATGTTCAACCTTCAAACTACAGTTTG 928
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 AACAAAGTTTGTAGGAATGTTACAAACACTAATGTCAACTAATGTCAACCTTCAAACTATAGTTTTG 780

Qy     929 TTGCTTTTAGTGCTGATGTAACAACCCCTCAATTATAAAATGCAAGAAGGACCGTTTGGA 988
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TTGCTTTTAGTGCTGATGTAACAACCTGTTAATTATTAATAAATGCAAGAAGGACCGTTTGAA 840

Qy     989 ATGGTGATGAACCTTCAAGTAGTAGAATTCTTGCAAACAACGAATAGTATCACAGATGTTTCTT 1048
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ATGGTGATGAACCTTCAAGTAGTAGAATTCTTGCAAACAACACCAATAGTATTACTGATGTTTCAT 900

Qy    1049 GGATTTATAGTTTAGTTGGAACAAACACGAAGTACCAATTTAGTTTTAGCAACTATGGTC 1108
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAATTTATAGTTTATTCTGGAACAAAACACGAAAATACCAATTTAGTTTTAGCAACTACGGTC 960

Qy    1109 CATCAACTGGTTATTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGCTAATAACG 1168
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CATCAACTGGTTATTTATATTTCCCTTATAGTTGGTTAAAGCGGCTGATGCGTAGTAATG 1020

Qy    1169 TTGGATTACAATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGTTGCCACTTCAA 1228
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TTGGATTACAATACAAATTAAATAATGGAAATGTTCAACCAGTTGAGTTTGCCACTTCAA 1080

Qy    1229 CTAGTGCAAAATAATACACAGCTAATCCAACTCCAGCAGTTGATGATGAGAAGTTGCTA 1288
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CTAGCGCAAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA 1140

Qy    1289 AAATCGTTTTATCAGGTT 1306
           ||||||||||||||||||
Db    1141 AAATCGTTTTATCAGGTT 1158
```

```
RESULT 6
ACF03364
ID   ACF03364 standard; DNA; 1152 BP.
XX
AC   ACF03364;
XX
DT   11-SEP-2003  (first entry)
XX
DE   Mycoplasma gallisepticum TTM-1 gene SEQ ID NO:1.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.
XX
PF   11-JUL-2002; 2002BP-00254879.
XX
PR   11-JUL-2001; 2001US-00901572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG ) ZEON CORP.
XX
PI   Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Claim 9; Page 30; 70pp; English.
XX
```

CC   The present invention describes a DNA molecule derived from a prokaryotic
CC   cell, where at least one of the DNA regions encoding NXB (where N =
CC   asparagine), X = any amino acid other than proline, and B = serine or
CC   threonine) has been modified so that no N-glycosylation occurs during the
CC   expression in a eukaryotic cell. Also described: (1) a fused DNA
CC   molecule, where a DNA encoding a signal sequence has been ligated to the
CC   N-terminal end of the modified DNA molecule as described above so that it
CC   may be expressed as a fusion protein; (2) a recombinant virus integrated
CC   with the DNA molecule or the fused DNA molecule described above; (3)
CC   producing a modified or fusion protein by using the recombinant virus
CC   described above, to express a protein encoded by the modified DNA
CC   molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC   vaccine comprising the recombinant virus. The DNA molecule has virucide
CC   and immunostimulant activities. The DNA molecule is useful for producing
CC   a vaccine for treating viral infections. The present sequence is used in
CC   the exemplification of the present invention
XX
SQ   Sequence 1152 BP; 437 A; 190 C; 188 G; 337 T; 0 U; 0 Other;

Query Match   78.7%; Score 1028; DB 7; Length 1152;
Best Local Similarity   100.0%; Pred. No. 1.2e-157;
Matches 1028; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   279 CTGTATGTCTATTACTAAAAAAAGATGCAAACCCAAATAATGCCAAACCCAATTAGAAGC 338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6 CTGTATGTCTATTACTAAAAAAAGATGCAAACCCAAATAATGCCAAACCCAATTAGAAGC 65

Qy   339 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 398
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACA 125

Qy   399 AGACTATGCCAAGATTGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 AGACTATGCCAAGATTGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAA 185

Qy   459 CAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGC 518
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 CAATAACCTTAATGCAACAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGC 245

Qy   519 CATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGA 578
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 CATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGA 305

Qy   579 AGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 AGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 365

Qy   639 ATCAACTGCTTATAATCAAATTCGCAATAATTCAAATTCGATCTATACAATAAAGCTAGTAG 698
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 ATCAACTGCTTATAATCAAATTCGCAATAATTCAAATTCGATCTATACAATAAAGCTAGTAG 425

Qy   699 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCTAATGAGAT 758
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 TTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTAGATTCTAATGAGAT 485

Qy   759 TACTACAGTTAATGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 TACTACAGTTAATGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAA 545

Qy   819 TGCTGATGCATTATCTATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTT 878
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546 TGCTGATGCATTATCTATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTT 605

Qy   879 TGTAGGGACTTTTACAAACGCTAATGTTCAACTTCAAACTACAGTTTTGCTTTTAG 938
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 TGTAGGGACTTTTACAAACGCTAATGTTCAACTTCAAACTACAGTTTTGCTTTTAG 665

Qy   939 TGCTGATGTAACACCCGTCAATTATAAATATCGAAGAAGGACCGTTGGAATGGTGATGA 998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 TGCTGATGTAACACCCGTCAATTATAAATATCGAAGAAGGACCGTTGGAATGGTGATGA 725

Qy   999 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAG 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAG 785

Qy   1059 TTTAGCTGGAACAAACAAACACGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGG 1118
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 TTTAGCTGGAACAAACAAACACGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGG 845

Qy   1119 TTATTTATATTCCCTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACA 1178
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 TTATTTATATTCCCTATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACA 905

Qy   1179 ATACAAATTAAATAATGGAAATGTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCAA 1238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 ATACAAATTAAATAATGGAAATGTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCAA 965

Qy   1239 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1025

Qy   1299 ATCAGGTT 1306
          ||||||||
Db   1026 ATCAGGTT 1033

RESULT 7
AAT96595
ID   AAT96595 standard; DNA; 1371 BP.
XX
AC   AAT96595;
XX
DT   15-JUL-1998 (first entry)
XX
DE   Hybrid Marek's disease virus/M. gallisepticum gene 40 K-S.
XX
KW   Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW   antigen; vaccine; poultry; ds.
XX
OS   Marek's disease gammaherpesvirus.
OS   Mycoplasma gallisepticum.
OS   Chimeric.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1371
FT   /*tag= a
FT   /product= "protein 40 K-S"
FT   misc_feature   1..192
FT   /*tag= b
FT   /note= "derived from Marek's disease virus gB gene"
FT   misc_feature   193..1368
FT   /*tag= c
FT   /note= "derived from gene encoding M. gallisepticum
FT   antigen"
XX
PN   WO9736924-A1.
XX
PD   09-OCT-1997.
XX
PF   28-MAR-1997; 97WO-JP001084.
XX
PR   29-MAR-1996; 96JP-00103548.
XX
PA   (JAPG) NIPPON ZEON KK.
XX
PI   Saito S, Tsuzaki Y, Yanagida N;
XX
DR   WPI; 1997-503046/46.
DR   P-PSDB; AAW36050.
XX
PT   Fusion protein comprising herpes virus outer membrane protein and
PT   antigenic polypeptide - for prevention of infection by Mycoplasma
PT   gallisepticum, especially in poultry.
XX
PS   Disclosure; Page 16-19; 51pp; Japanese.
XX

```
CC   This sequence encodes the chimeric protein 40 K-S which comprises a
CC   fragment of the Marek's disease virus outer membrane protein gB fused to
CC   an antigenic protein from Mycoplasma gallisepticum. The chimeric protein
CC   can be used in recombinant live vaccines for prevention of infection by
CC   Mycoplasma gallisepticum, especially as the outer membrane protein shows
CC   antigenicity in poultry
XX
SQ   Sequence 1371 BP; 491 A; 242 C; 232 G; 406 T; 0 U; 0 Other;

Query Match    78.4%;  Score 1024.2;  DB 2;  Length 1371;
Best Local Similarity  99.7%;  Pred. No. 5.1e-157;
Matches 1026;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   278 GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 337
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 247

Qy   338 CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 397
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 307

Qy   398 AAGACTATGCCAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA 457
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 AAGACTATGCCAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA 367

Qy   458 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAG 517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAG 427

Qy   518 CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG 577
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   428 CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG 487

Qy   578 AAGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGT 637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 AAGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGT 547

Qy   638 CATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTA 697
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 CATCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTA 607

Qy   698 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 667

Qy   758 TTACTACAGTTAATCGGAATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 817
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 TTACTACAGTTAATAGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 727

Qy   818 ATGCTGATGCATTATCTAATAGTTTTATTATTAAAAAGTGATTCAAAATAATGAACAAAAGTT 877
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 ATGCTGATGCATTATCTAATAGTTTTATTATTAAAAAGTGATTCAAAATAATGAACAAAAGTT 787

Qy   878 TTGTAGGGACTTTACAAACGCTAATGTTCAAACTTCAAACTACAGTTTTGTTGCTTTTA 937
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 TTGTAGGGACTTTACAAACGCTAATGTTCAAACTTCAAACTACAGTTTTGTTGCTTTTA 847

Qy   938 GTGCTGATGTAACACCCGTCAATTATAAAATGCAAGAAGGACCCGTTGGAATGGTGATG 997
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 GTGCTGATGTAACACCCGTCAATTATAAAATGCAAGAAGGACCCGTTGGAATGGTGATG 907

Qy   998 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCCTTGGATTTATA 1057
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTCCTTGGATTTATA 967

Qy   1058 GTTTAGCTGCAACAAACACCGAAGTACCAATTTAGTTTAGCAACTATGTGCCATCAACTG 1117
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GTTTAGCTGGAACAAACACCGAAGTACCAATTTAGTTTAGCAACTATGGTCCATCAACTG 1027

Qy   1118 GTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTAATAACGTTGGATTAC 1177
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 GTTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTAATAACGTTGGATTAC 1087
```

```
Qy   1178 AATACAAATTAAATAATATGGAAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 1237
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 AATACAAATTAAATAATATGGAAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 1147

Qy   1238 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTT 1297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTT 1207

Qy   1298 TATCAGGTT 1306
              |||||||||
Db   1208 TATCAGGTT 1216
```

```
RESULT 8
AAT96596
ID   AAT96596 standard; DNA; 3261 BP.
XX
AC   AAT96596;
XX
DT   15-JUL-1998  (first entry)
XX
DE   Hybrid Marek's disease virus/M. gallisepticum gene 40 K-C.
XX
KW   Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW   antigen; vaccine; poultry; ds.
XX
OS   Marek's disease gammaherpesvirus.
OS   Mycoplasma gallisepticum.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3261
FT                   /*tag= a
FT                   /product= "protein 40 K-C"
FT   misc_feature    1..2016
FT                   /*tag= b
FT                   /note= "derived from Marek's disease virus gB gene"
FT   misc_feature    2077..3258
FT                   /*tag= c
FT                   /note= "derived from gene encoding M. gallisepticum
FT                   antigen"
XX
PN   WO9736924-A1.
XX
PD   09-OCT-1997.
XX
PF   28-MAR-1997;  97WO-JP001084.
XX
PR   29-MAR-1996;  96JP-00103548.
XX
PA   (JAPG ) NIPPON ZEON KK.
XX
PI   Saito S, Tsuzaki Y, Yanagida N;
XX
DR   WPI; 1997-503046/46.
DR   P-PSDB; AAW36051.
XX
PT   Fusion protein comprising herpes virus outer membrane protein and
PT   antigenic polypeptide - for prevention of infection by Mycoplasma
PT   gallisepticum, especially in poultry.
XX
PS   Disclosure; Page 22-30; 51pp; Japanese.
XX
CC   This sequence encodes the chimeric protein 40 K-C which comprises a
CC   fragment of the Marek's disease virus outer membrane protein gB fused to
CC   an antigenic protein from Mycoplasma gallisepticum. The chimeric protein
CC   can be used in recombinant live vaccines for prevention of infection by
CC   Mycoplasma gallisepticum, especially as the outer membrane protein shows
CC   antigenicity in poultry
XX
SQ   Sequence 3261 BP; 1099 A; 617 C; 648 G; 897 T; 0 U; 0 Other;

Query Match    78.3%;  Score 1022.6;  DB 2;  Length 3261;
```

Best Local Similarity   99.6%;   Pred. No. 9e-157;
Matches 1025;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

```
Qy    278  GCTGTATGTCTATTACTAAAAAAAGATGCAAACCCAATTAATGCCAAACCCAATTAGAAG  337
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2078  GCTGTATGTCTATTACTAAAAAAAGATGCAAACCCAATTAATGCCAAACCCAATTAGAAG 2137

Qy    338  CAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC  397
           |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2138  CAGCCCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 2197

Qy    398  AAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAAACAGTTA  457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2198  AAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAAACAGTTA 2257

Qy    458  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAG  517
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2258  ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAG 2317

Qy    518  CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG  577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2318  CCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG 2377

Qy    578  AAGCATACAAAGACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT  637
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2378  AAGCATACAAAGACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT 2437

Qy    638  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATACAATAAAGCTAGTA  697
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2438  CATCAACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATACAATAAAGCTAGTA 2497

Qy    698  GTTTAATAACTAAAAACACTAGTATCCACTAAAATGGGGAACGCTTTAGATTCTAATGAGA  757
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2498  GTTTAATAACTAAAAACACTAGTATCCACTAAAATGGGGAACGCTTTAGATTCTAATGAGA 2557

Qy    758  TTACTACAGTTAATCGGAATATAATAATACGTTATCAACTATTAATGAACAAAAGACTA  817
           ||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||||
Db   2558  TTACTACAGCTAATAAGAATATATAATAATACGTTATCAACTATTAATGAACAAAAGACTA 2617

Qy    818  ATGCTGATGCAATATCTAATAGTTTTATTAATAAAAGTGATTCAAAATAATGAACAAAGTT  877
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   2618  ATGCTGATGCAATATCTAATAGTTTTATTAATAAAAGTGATTCAAAATAATGAACAAAGTT 2677

Qy    878  TTGTAGGGACTTTTACAAACGCTAATGTCAACCTTCAAACTACAGTTTGTTGCTTTTA  937
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2678  TTGTAGGGACTTTTACAAACGCTAATGTCAACCTTCAAACTACAGTTTGTTGCTTTTA 2737

Qy    938  GTGCTGATGTAACACCCGTCAATTATAAAATGCAAGAAGGAACCGTTTGGAATGGTGATG  997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2738  GTGCTGATGTAACACCCGTCAATTATAAAATGCAAGAAGGAACCGTTTGGAATGGTGATG 2797

Qy    998  AACCTTCAAGTAGATAGAAATTCTTGCAAACACGAGATAGTATCACAGATGTTTCTTGATTATA 1057
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2798  AACCTTCAAGTAGATAGAAATTCTTGCAAACACGAGATAGTATCACAGATGTTTCTTGATTATA 2857

Qy   1058  GTTTAGCTGAACAAACACGAGTACCAATTTAGTTTTAGCACTATGGTCCATCAACTG 1117
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2858  GTTTAGCTGAACAAACACGAGTACCAATTTAGTTTTAGCACTATGGTCCATCAACTG 2917

Qy   1118  GTTATTTATAATTCCCTTATAAGTTGGTAAAGCAGCTGAATGCTAATAACGTTGGATTAC 1177
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2918  GTTATTTATAATTCCCTTATAAGTTGGTAAAGCAGCTGAATGCTAATAACGTTGGATTAC 2977

Qy   1178  AATACAAAATTAAATAAATAAGTGAAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAA 1237
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2978  AATACAAAATTAAATAAATAAGTGAAAATGTTCAACAAGTTGCCACTTCAACTAGTGCAA 3037

Qy   1238  ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1297
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3038  ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 3097

Qy   1298  TATCAGGTT 1306
           |||||||||
Db   3098  TATCAGGTT 3106
```

RESULT 9
AAT04076
ID   AAT04076   standard; DNA; 1015 BP.
XX
AC   AAT04076;
XX
DT   19-JUL-1996 (first entry)
XX
DE   M.gallisepticum DNA sequence II encodes 261 amino acid protein.
XX
KW   Detection; probe; primer; PCR; amplification; secretion; lung;
KW   avian chronic respiratory disease; respiratory tract; nasal cavity; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key   Location/Qualifiers
FT   CDS   202..987
FT   /*tag= a
XX
PN   JP07236498-A.
XX
PD   12-SEP-1995.
XX
PF   25-FEB-1994;   94JP-00052764.
XX
PR   25-FEB-1994;   94JP-00052764.
XX
PA   (JAPG) NIPPON ZEON KK.
PA   (SHIO) SHIONOGI & CO LTD.
XX
DR   WPI; 1995-347462/45.
DR   P-PSDB; AAR79911.
XX
PT   Detection of Mycoplasma gallisepticum - for the quick detection, i.e.
PT   within one day, of avian chronic respiratory diseases.
XX
PS   Claim 3; Page 10-11; 11pp; Japanese.
XX
CC   This is the nucleotide sequence of a fragment of the Mycoplasma
CC   gallisepticum genome which codes for a 261 amino acid protein. This
CC   sequence and the sequence of AAT04075 (encoding a 661 amino acid protein)
CC   can be used to detect M.gallisepticum using probes based on nucleotides
CC   1125-1648 and primers based on nucleotides 449-466, the complement of
CC   bases 893-919, 1908-1934 and the complement of bases 2184-2210 of
CC   AAT04075, and a probe based on nucleotides 718-41 of this sequence. The
CC   method using these sequences is faster i.e. is able to detect
CC   M.gallisepticum, which causes avian chronic respiratory diseases, within
CC   one day, from avian secretions, washings from the lung, respiratory
CC   tract, nasal cavity, etc
XX
SQ   Sequence 1015 BP; 400 A; 162 C; 138 G; 315 T; 0 U; 0 Other;

Query Match   75.8%;   Score 989.4;   DB 2;   Length 1015;
Best Local Similarity 98.4%;   Pred. No. 2.2e-151;
Matches 999;   Conservative   0;   Mismatches   16;   Indels   0;   Gaps   0;

```
Qy      1  AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAAATCTTCTAACAA   60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AAAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAAAACACAAAAATCTTCTAACAA   60

Qy     61  AATCCTAAATAAATAAGCCGTTAAATAACTAAAAAATTAAAAAAATGGTTTTTCTTATC  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AATCCTAAATAAATAAGCCGTTAAATAACTAAAAAATTAAAAAAATGGTTTTTCTTATC  120

Qy    121  AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGAT  180

Qy    181  ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db	ATAAATATATCTTAATATTCTATGAATAGAAAAGAATCATCTTAAAGACATTAGTTTG	240
Qy	TTAGGTACAACATCCTTTCTTAGCATTGGGATTCTAGCTGTATGTCGTATTACTAAAAA	300
Db	TTAGGTACAACATCCTTTCTTAGCATTGGGATTCTAGCTGTATGTCTATTACTAAAAA	300
Qy	GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT	360
Db	GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCAGCGCGAATGGAGTTAACTGAT	360
Qy	CTAATCAATGCTAAAGCCGATGACATTAGTTCACTACAAGACTATGCCAAGATTGAAGCT	420
Db	CTAATCAATGCTAAAGCCAAGGACACATTAGCTTCACTACAAGACTATGCTAAGATTGAAGCT	420
Qy	AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA	480
Db	AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACACTA	480
Qy	GAACAACTAAAAATGGCTAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT	540
Db	GAACAACTAAAAATGGCTAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT	540
Qy	AAAACGACTTTGATAATGAACAACCCAAATTTAGTTGAAGCATACAAGAGCACTAAAAACC	600
Db	AAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCATACAAGAGCACTAAAAACC	600
Qy	ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT	660
Db	ACTTTAGAACAACGTGCTACTAACCCTTGAAGGTTTAGCTTCAACTGCTTATAATCAGATT	660
Qy	CGCAATAATTAGTTGGATCTATACAATAAGCTAGTAGTTTATAATAACTAAAACACTAGAT	720
Db	CGTAATAATTAGTTGGATCTATACAATAATGCTAGTAGTTTATAATAACTAAAACACTAGAT	720
Qy	CCACTAAATGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT	780
Db	CCACTAAATGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT	780
Qy	AATAATACGTTATCAACCATTAATGAACAAAGAACTAATGCTGATGCATTATCTAATAGT	840
Db	AATAATACGTTATCAACCATTAATGAACAAAGAACTAATGCTGATGCATTATCTAATAGT	840
Qy	TTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTGAGGGACTTTGTAGGGGACTTTACAAACGCT	900
Db	TTTATTAAAAAGTGATTCAAAATAATGAACAAGTTTGTGAGGGACTTTGTAGGGGACTTTACAAACGCT	900
Qy	AATGTTCAACCTTCAAACTACAGTTTTGTGCTTTAGTGCTGATGTAACACCCGTCAAT	960
Db	AATGTTCAACCTTCAAACTACAGTTTTGTGCTTTAGTGCTGATGTAACACCCGTCAAT	960
Qy	TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC	1015
Db	TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTC	1015

RESULT 10
ACF03382
ID   ACF03382   standard; DNA; 1082 BP.
XX
AC   ACF03382;
XX
DT   11-SEP-2003   (first entry)
XX
DE   M. gallisepticum TTM-1 portion of modified pNZ40K-S gene SEQ ID NO:24.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.

PF   11-JUL-2002; 2002EP-00254879.
XX
PR   11-JUL-2001; 2001US-00901572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG) ZEON CORP.
XX
PI   Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
DR   P-PSDB; ABR57374.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Example 1; Page 42-43; 70pp; English.
XX
CC   The present invention describes a DNA molecule derived from a prokaryotic
CC   cell, where at least one of the DNA regions encoding NXB (where N =
CC   asparagine, X = any amino acid other than proline, and B = serine or
CC   threonine) has been modified so that no N-glycosylation occurs during the
CC   expression in a eukaryotic cell. Also described: (1) a fused DNA
CC   molecule, where a DNA encoding a signal sequence has been ligated to the
CC   N-terminal end of the modified DNA molecule as described above so that it
CC   may be expressed as a fusion protein; (2) a recombinant virus integrated
CC   with the DNA molecule or the fused DNA molecule described above; (3)
CC   producing a modified or fusion protein by using the recombinant virus
CC   described above, to express a protein encoded by the modified DNA
CC   molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC   vaccine comprising the recombinant virus. The DNA molecule has virucide
CC   and immunostimulant activities. The DNA molecule is useful for producing
CC   a vaccine for treating viral infections. The present sequence is used in
CC   the exemplification of the present invention
XX
SQ   Sequence 1082 BP; 403 A; 184 C; 183 G; 312 T; 0 U; 0 Other;

Query Match   70.6%; Score 921.6; DB 7; Length 1082;
Best Local Similarity   98.0%; Pred. No. 2e-140;
Matches   933;   Conservative   0; Mismatches 19; Indels   0; Gaps   0;

Qy	355	ACAGAGATCTAATCAATGCTAAAGCCGATGACATTAGTTCACTACAAGACTATGCCAAGATT	414
Db	1	ACAGAGATCTAATCAATGCTAAAGCCGATGACATTAGTTCACTACAAGACTATGCCAAGATT	60
Qy	415	GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA	474
Db	61	GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA	120
Qy	475	ACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT	534
Db	121	ACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT	180
Qy	535	ACGGATAAAACGACTTTGATAATGAACACCCCAAATTTAGTTGAAGCATACAAAGCACTA	594
Db	181	ACGGATAAAACGACTTTGATAATGAACACCCCAAATTTAGTTGAAGCATACAAAGCACTA	240
Qy	595	AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT	654
Db	241	AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAAT	300
Qy	655	CAAATTCGCAATAATTCGATTATCAATGAAGCTAGTAGTTTAATAACTAAAACA	714
Db	301	CAAATTCGCAATAATTCGATTATCAATGAAGCTAGTAGTTTAATAACTAAAACA	360
Qy	715	CTAGATCCACTAAATGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGG	774
Db	361	CTAGATCCACTAAATGGGGAACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATAAG	420
Qy	775	AATATTAATAACGTTATCAACTATTAATGAACAAAAGAACTAATGCTGATGCATTATCT	834
Db	421	AATATCCAGAATACGTTATCAACTATTAATGAACAAAAGAACTAATGCTGATGCATTATCT	480

```
CC   DNA to produce fusion proteins with an N-terminal bacterial enzyme
CC   sequence. See also AAQ04686 and AAQ05649-53. (Updated on 10-MAR-2003 to
XX   add missing Os field.)
SQ   Sequence 853 BP; 329 A; 138 C; 128 G; 258 T; 0 U; 0 Other;

Query Match         63.1%;  Score 824.2; DB 1;  Length 853;
Best Local Similarity 97.9%;  Pred. No. 1.2e-124;
Matches  835;  Conservative  0;  Mismatches  18;  Indels  0;  Gaps  0;

Qy    163 TTAGTCATCTTTAAGATATAAATATCTTAATATTCTATGAATAAGAAAAGAATCATC 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTAGTCATCTTTAAGATATAAATATCTTAATATTCTATGAATAAGAAAAGAATCATC  60

Qy    223 TTAAAGACTATTAGTTTGTTAGTAGTACAACATCCTTCTTAGCATTGGGATTCTCAGCTGT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTAAAGACTATTAGTTTGTTAGTAGTACAACATCCTTCTTAGCATTGGGATTCTCAGCTGT 120

Qy    283 ATGTCTATTACTAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCG 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATGTCTATTACTAAAAAGACGCAAACCCAATAATGGCCAAACCCAATTACAAGCAGCG 180

Qy    343 CGAATGGAGTTAACAGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGAC 402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGAATGGAGTTAACTGATCTAATCAATGCTAAAGGCGATGACATTAGCTTCACTACAAGAC 240

Qy    403 TATGCCAAGATGAAGCTAGTTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAACAAT 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATGCTAAGATGAAGCTAGTTTATCATCGTGCTTATAGTGAAGCTGAAACAGTTAACAAT 300

Qy    463 AACCTTAATGCAACATTTGAACAACAACAACTAAATGGCTAAAACTAATTTAGAATCAGCCATC 522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AACCTTAATGCAACACTAGAACAACAACAACTAAATGGCTAAAACTAATTTAGAATCAGCCATC 360

Qy    523 AACCAAGCTAATACGGATAATAAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCA 582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AACCAAGCTAATACGGATAATAAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCA 420

Qy    583 TACAAAGCACTAAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCA 642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TACAAAGCACTAAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTAGCTTCA 480

Qy    643 ACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGTACTAGTAGTTTA 702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACTGCTTATAATCAGATTCGTAATAATTAGTGGATCTATACAATAATGCTAGTAGTTTA 540

Qy    703 ATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTGAGATTCTAATGAGATTACT 762
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ATAACTAAAACACTAGATCCACTAAATGGGGAATGCTTTTGAGATTCTAATGAGATTACT 600

Qy    763 ACAGTTAATCGGAATTAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCT 822
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACAGTTAATCGGAATTAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCT 660

Qy    823 GATGCATTATTATTAATAGTTTTATTAATAGTTTTTATTAAAAATGATTCAAAAATAATGAACAAAGTTTTGTA 882
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GATGCATTATTATTAATAGTTTTATTAATAGTTTTTATTAAAAATGATTCAAAAATAATGAACAAAGTTTTGTA 720

Qy    883 GGGACTTTACAAACGTAATGCTAATGTTCAACCTTCAAACTACAGTTTTGTTTGTTAGTGCT 942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGACTTTACAAACGTAATGCTAATGTTCAACCTTCAAACTACAGTTTTGTTTGTTAGTGCT 780

Qy    943 GATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGGTTTGATGAACCT 1002
          |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    781 GATGTAACACCCGTCAATTATAAATATGCAAGAAGAGACGGTTGAAATGGTGATGAACCT 840

Qy   1003 TCAAGTAGAATTC 1015
          ||||||||||||
Db    841 TCAAGTAGAATTC 853
```

```
Qy    835 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA 894
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA 540

Qy    895 AACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTTAGTGTCTGATGTAACACCC 954
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AACGCTAATGTTCAACCTAGCCAGTACAGTTTTGTTGCTTTAGTGTCTGATGTAACACCC 600

Qy    955 GTCAATTATAAAATGCAAGAAGGAACCGTTGAATGGTGATGAACCTTCAAGTAGAATT 1014
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTCAATTATAAAATGCAAGAAGGAACCGTTGAATGGTGATGAACCTTCAAGTAGAATT 660

Qy   1015 CTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAAC 1074
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTTGCAAACACGAATAGTATCACAGATGTTCTTGGATTTATAGTTTAGCTGGAACAAAC 720

Qy   1075 ACGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCT 1134
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACGAAGTACCAATTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCT 780

Qy   1135 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAAT 1194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATAAGTTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAAT 840

Qy   1195 GGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATAATACTACAGCTAAT 1254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GGAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGCGCAGCAGACTACAGCTAAT 900

Qy   1255 CCAACTCCAGCAGTTGATGAGATTAAAGTGCTAAAATCGTTTTATCAGGTT 1306
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCAACTCCAGCAGTTGATGAGATTAAAGTGCTAAAATCGTTTTATCAGGTT 952
```

```
RESULT 11
AAQ04687
ID   AAQ04687 standard; DNA; 853 BP.
XX
AC   AAQ04687;
XX
DT   10-MAR-2003 (revised)
DT   08-OCT-1990 (first entry)
XX
DE   DNA encoding TMG-1 antigen.
XX
KW   Mycoplasma gallisepticum; poultry; vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
PN   JP02111795-A.
XX
PD   24-APR-1990.
XX
PP   02-JUN-1989;  89JP-00140283.
PR   02-JUN-1988;  88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Kodama K, Saito S, Yanagida N, Kanogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAR05082.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Claim 5; Fig 2; 20pp; Japanese.
XX
CC   The DNA can be inserted into an expression vector for the prodn. of
CC   polypeptide which elicits an antigen-antibody reaction with anti-
CC   mycoplasma gallisepticum poultry sera. It may also be ligated to other
```

```
RESULT 12
AAN92574
```

```
ID   AAN92574 standard; DNA; 853 BP.
XX
AC   AAN92574;
XX
DT   25-MAR-2003 (revised)
DT   11-MAY-1990 (first entry)
XX
DE   DNA sequence of TM-1 encoding polypeptide TMG-1.
XX
KW   Mycoplasma gallisepticum; Poultry vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key            Location/Qualifiers
FT   CDS            40..822
FT                  /*tag= a
XX
PN   EP345021-A.
XX
PD   06-DEC-1989.
XX
PF   31-MAY-1989;   89EP-00305441.
XX
PR   02-JUN-1988;   88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI SEIYAKU KK.
XX
PI   Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAP93959.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Disclosure; Fig 2; 31pp; English.
XX
CC   This base sequence of TM-1 encodes the TMG-1 polypeptide which has the
CC   same sequence as that of a polypeptide expressed in Mycoplasma
CC   gallisepticum in nature. When inserted into a recombinant vector used to
CC   transform a host the antigen protein produced can be used as a vaccine to
CC   prevent and diagnose MG infection. (Updated on 25-MAR-2003 to correct PF
CC   field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
CC   2003 to correct PI field.)
XX
SQ   Sequence 853 BP; 330 A; 139 C; 128 G; 256 T; 0 U; 0 Other;

Query Match             61.4%;  Score 801.8;  DB 1;  Length 853;
Best Local Similarity   96.2%;  Pred. No. 5e-121;
Matches  821;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy    163 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTTATGATAAGAGAAAGAATCATC 222
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1 TTAGTCATCTTTAAGATATAAATATATCTTAATATTCTTATGATAAGAGAAAGAATCATC 60

Qy    223 TTAAAGACTATTAGTTGTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGT 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTAAAGACTATTAGTTGTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGT 120

Qy    283 ATGTCTATTACTAACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCG 342
          ||||||||||||||||||||||| |||||||||||||||||||||||||||||||| |||||
Db    121 ATGTCTATTACTAACTAAAAAGACGCAAACCCAAATAATGGCCAAACCCAATTACAGCAGCG 180

Qy    343 CGAATGAGTAACAGATCTAATCAATGCTAAAGCTAAAGCAGGACATTAGCTTCACTACAAGAC 402
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGAATGAGTAACTGATCTAATCAATGCTAAAGCTAAAGCAGGACATTAGCTTCACTACAAGAC 240

Qy    403 TATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 462
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATGCTAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAAT 300

Qy    463 AACCTTAATGCAACATTAGAACAACTAAAAAATGGCTAAAACTAATTAGAATCAGCCATC 522
```

```
Db    301 AACCTTAATGCAACACTAGAACACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATC 360
          ||||||||||||||| |||||||| ||||||||||||||||||| |||||||||||||||
Qy    523 AACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCA 582
          |||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||||
Db    361 AACCAAGCTAATACGGATAAAACGACTTTGATAATGAACATCCAAATTTAGTTGAAGCA 420

Qy    583 TACAAAGACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCA 642
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||
Db    421 TACAAAGACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCA 480

Qy    643 ACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTA 702
          ||||||||||||||| ||||| ||||||||||||||||||||||| ||||||||||||||
Db    481 ACTGCTTATAATCAGATTCGTAATAATTAGTGGATCTATACAATAACAATAATGCTAGTAGTTTA 540

Qy    703 ATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGAATTCTAATGAGAGATTACT 762
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    541 ATAACTAAAAACACTAGATCCACTAAATGGGGAATGCTTTTAGAATTCTAATGAGAGATTACT 600

Qy    763 ACAGTTAATCGGAATATTAATAATACGTTATCAACCATTAATGAACAAAAGACTAATGCT 822
          |||||||||||||||||| |||||||||||||| ||||||||||||||| |||||||||||
Db    601 ACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAACAAAGACTAATGCT 660

Qy    823 GATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTA 882
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GATGCATTATCTAATAGTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTA 720

Qy    883 GGGACTTTTACAAAGCTAATGTTCAACCTCAAACTACAGTTTGTTTGCTTTTAGTGCT 942
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGGACTTTTACAAACGCTAATGTTCAACCTCAAACTACAGTTTGTTTGCTTTTAGTGCT 780

Qy    943 GATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCT 1002
          |||||||||||||||||||||||||||| ||||| ||| |||||| |||||||||||||||
Db    781 GATGTAACACCCGTCAATTATAAATATACAAGAAGAACGGTTTGAAATGGTGATGAACCT 840

Qy   1003 TCAAGTAGAATTC 1015
          |||||||||||||
Db    841 TCAAGTAGAATTC 853
```

```
RESULT 13
AAN92568
ID   AAN92568 standard; DNA; 708 BP.
XX
AC   AAN92568;
XX
DT   25-MAR-2003 (revised)
DT   11-MAY-1990 (first entry)
XX
DE   DNA sequence of M1 encoding Mycoplasma gallisepticum (MG-1) polypeptide.
XX
KW   Mycoplasma gallisepticum; Poultry vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP345021-A.
XX
PD   06-DEC-1989.
XX
PF   31-MAY-1989;   89EP-00305441.
XX
PR   02-JUN-1988;   88JP-00136343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI SEIYAKU KK.
XX
PI   Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAP93646.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
```

Left column

```
XX
PS   Disclosure; Fig 1a; 31pp; English.
XX
CC   This base sequence of M1 encodes the MG1 polypeptide which elicits an
CC   antigen-antibody reaction with anti-MG poultry sera. When inserted into a
CC   recombinant vector used to transform a host the antigen protein produced
CC   can be used as a vaccine to prevent and diagnose MG infection. (Updated
CC   on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct
CC   PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ   Sequence 708 BP; 278 A; 119 C; 107 G; 204 T; 0 U; 0 Other;

Query Match         52.1%;  Score 679.8; DB 1;  Length 708;
Best Local Similarity   97.6%;  Pred. No. 2.7e-101;
Matches  690;  Conservative   0;  Mismatches  17;  Indels   0;  Gaps   0;

Qy    280 TGTATGTCTATTACTAATAAAAAGATGCAAACCCAATTAATGGCCAATTAGAAGCA 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTATGTCTATTACTAATAAAAAGACGCAAACCCAAATAATGGCCAATTACAAGCA 60

Qy    340 GCGCGAATGGAGTTAACAGATCTAATGTCTAAAGCGATGACATTAGCTTCACTACAA 399
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     61 GCGCGAATGGAGTTAACTGATCTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAA 120

Qy    400 GACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 459
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GACTATGCTAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 180

Qy    460 AATAACCTTAATGCAACATTAGAACACGTGCTAAAAATGGCTAAAACTAATTTAGAATCAGCC 519
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATAACCTTAATGCAACACTAGAACAACAACTAATTTAGAATCAGCC 240

Qy    520 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCCAAATTAGTTGAA 579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACATCCCAAATTTAGTTGAA 300

Qy    580 GCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGTCA 639
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTAGCT 360

Qy    640 TCAACTGCTTATTAATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGT 699
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAACTGCTTATTATCAGATTCGTAATAATTTAGTGGATCTATACAATAAATGCTAGTAGT 420

Qy    700 TTAATAACTAAAAACACTAGATCCACTAAATGGGGGGAACGCTTTTAGATTCTAATGAGATT 759
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTAATAACTAAAAACACTAGATCCACTAAATGGGGGGAATGCTTTTAGATTCTCAATGAGATT 480

Qy    760 ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAAT 819
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAAT 540

Qy    820 GCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTT 879
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCTGATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTT 600

Qy    880 GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGT 939
          |||||||||||||||||||||||||||||||||||||||||
Db    601 GTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGT 660

Qy    940 GCTGATGTAACACCCCGTCAATTATAAATATGCAAGAAGGACCGTTTG 986
          ||||||||||||||||||||||||||| |||||
Db    661 GCTGATGTAACACCCCGTCAATTATAAATATGCAAGAAGGACCGTTTG 707

RESULT 14
AAQ04686
ID   AAQ04686 standard; DNA; 708 BP.
XX
AC   AAQ04686;
XX
DT   10-MAR-2003 (revised)
DT   08-OCT-1990 (first entry)
```

Right column

```
XX
DE   DNA encoding MG-1 antigen.
XX
KW   Mycoplasma gallisepticum; poultry; vaccine; ss.
XX
OS   Mycoplasma gallisepticum.
XX
PN   JP02111795-A.
XX
PD   24-APR-1990.
XX
PF   02-JUN-1989;    89JP-00140283.
XX
PR   02-JUN-1988;    88JP-00135343.
XX
PA   (JAPG ) NIPPON ZEON KK.
PA   (SHIO ) SHIONOGI & CO LTD.
XX
PI   Kodama K,  Saito S,  Yanagida N,  Kamogawa K,  Iritani Y,  Aoyma S;
XX
DR   WPI; 1989-358393/49.
DR   P-PSDB; AAR05081.
XX
PT   New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT   vaccines.
XX
PS   Claim 3; Fig 1a; 20pp; Japanese.
XX
CC   The DNA can be inserted into an expression vector for the prodn. of
CC   polypeptide which elicits an antigen-antibody reaction with anti-
CC   mycoplasma gallisepticum poultry sera. It may also be ligated to other
CC   DNA to produce fusion proteins with an N-terminal bacterial enzyme
CC   sequence. See also AAQ04687 and AAQ05649-53. (Updated on 10-MAR-2003 to
CC   add missing OS field.)
XX
SQ   Sequence 708 BP; 279 A; 119 C; 107 G; 203 T; 0 U; 0 Other;

Query Match          51.9%;  Score 678.2; DB 1;  Length 708;
Best Local Similarity   97.5%;  Pred. No. 5e-101;
Matches  689;  Conservative   0;  Mismatches  18;  Indels   0;  Gaps   0;

Qy    280 TGTATGTCTATTACTAATAAAAAGATGCAAACCCAATTAATGGCCAATTAGAAGCA 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTATGTCTATTACTAATAAAAAGACGCAAACCCAAATAATGGCCAATTACAAGCA 60

Qy    340 GCGCGAATGGAGTTAACAGATCTAATGTCTAAAGCGATGACATTAGCTTCACTACAA 399
          |||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     61 GCGCGAATGGAGTTAACTGATCTAATCAATGCTAAAGCAAGGACATTAGCTTCACTACAA 120

Qy    400 GACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 459
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GACTATGCTAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAAC 180

Qy    460 AATAACCTTAATGCAACATTAGAACACGTGCTAAAAATGGCTAAAACTAATTTAGAATCAGCC 519
          ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    181 AATAACCTTAATGCAACACTAGAACAACAACTAATTTAGAATCAGCC 240

Qy    520 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCCAAATTAGTTGAA 579
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACATCCCAAATTTAGTTGAA 300

Qy    580 GCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGTCA 639
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTAGCT 360

Qy    640 TCAACTGCTTATAATCAAATTCGCAATAATTTAGTGGATCTATACAATAAAGCTAGTAGT 699
          |||||||||||  |||||||||||| |||||||||||||||||||||||||  ||||||
Db    361 TCAACTGCTTATAATCAGATTCGTAATAATTTAGTGGATCTATACAATAAATGCTAGTAGT 420

Qy    700 TTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATT 759
          |||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    421 TTAATAACAAAACACTAGATCCACTAAATGGGGGAATGCTTTTAGATTCTCAATGAGATT 480
```

```
CC   sequence, and a probe based on nucleotides 718-41 of AAT04076. The method
CC   using these sequences is faster i.e. is able to detect M.gallisepticum,
CC   which causes avian chronic respiratory diseases, within one day, from
CC   avian secretions, washings from the lung, respiratory tract, nasal
XX   cavity, etc
SQ   Sequence 2196 BP; 800 A; 353 C; 376 G; 667 T; 0 U; 0 Other;

Query Match             16.6%;  Score 216.6;  DB 2;  Length 2196;
Best Local Similarity   55.1%;  Pred. No. 2.3e-26;
Matches  548;  Conservative  0;  Mismatches  414;  Indels  33;  Gaps  5;

Qy    329 AATTAGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAG   388
          ||||||  |||   | | |||||||||| | ||  ||| |||||||||   |||  ||||
Db    352 AATTAGCAGACGCAAAAGCTGCTTTAACTACTTTGATTAATGGTGAAACTGCAAATCTTG   411

Qy    389 CTTCACTACAGAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  448
          | ||||  | |||||||||| | |||| ||| ||| |||| ||||| | ||||| |||||
Db    412 CGTCATATGAAGAGACTATGTCTAGATCAAGAAGTGAATTAACATCAGCGTATGAACAGCTA 471

Qy    449 AAACAGTTAACAATAACCTTAATGCAACCAAGCTAATACGAGATCACAT            508
          || |||||  |  ||    | || | | | ||    ||||| ||| ||||
Db    472 AAGCAGTTGCAGTCAGTAAATCGGTGCAACTCTAAATCGAGTTAATACACAT         531

Qy    509 TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAA   568
          |||||||||  ||||  ||||| | || ||| |||| || |  ||||||| |||  ||
Db    532 TAGATGCTGCTATTAAAAAGCGTCAGTGCTAAGAATGATTTTGATGCACAGCACGGGGT    591

Qy    569 ATTTAGTGAAGCATACAAAGCACCACTTAAAAACCACTTAAGAAACGTAAAAGAAG...   628
          |||||||||| ||| || || || ||||  ||| ||||| |    ||| ||| | ||
Db    592 CACTAGTGGAGCATATCAATCTAAAGAAACGTAAAGAGAAGAAAAAACTAATTTAG      651

Qy    629 AAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATA   688
          ||| |||  | || | | |||  || | || | || ||||  | ||  |  ||||
Db    652 ATTCTCTGCAAACGAAAATTATGCAGCAATCAGTAATCTTATAGTTTATTATATGAAA    711

Qy    689 AAGCTAGTAGTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATT   748
          || |  |||||  |||| | | | || | || ||| | ||  | |||| | || |||
Db    712 AAGCCAATACATTGTTACAGTGTTACAGCCCTGCTACTGGAAATATTCCTGAAGTTA    771

Qy    749 CTAATGAATTACTACAGTTAATCGGAATAATAATACGTTATCAACTATTAATGAAC   808
          || |||  |  | | | |  |||   | |||| |||| || |||| | ||||  | |
Db    772 TGAGTG---TAACAGTCGTAATCAAGATATTACTACTAGTTTTATTAAAAAGTGAT     828

Qy    809 AAAGACTAATGCTAATGCATTATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATG   868
          || | | || |   ||| | | | |||   |  | | |   | ||  |   |||| |
Db    829 GAAACAAAATGCTGATAAATCTGATAATTTAGCTAACAGTTTTATCARACAGTCTTTAGTTAAAAAATA 888

Qy    869 AACAAAGTTTTGTAGGGACTTTTACAAACGCTAATGTTCAAACCTAATGTTCAAACCAGTTTTG  928
          ||| |||||||| ||  |    |   | |  | | | | | | ||||   | |||   ||||||
Db    889 AATTGACTAGAGTTG---ATGTAGCAAATAATCAGGAGCAACCAGCAAATTACAGTTTTG       945

Qy    929 TTGCTTTTAGTGCTGATGCTGATAGTGAACACCCGTCAATTATAAATATGCAAGAAGACCGTTTGGA 988
          || |||  | | | ||| ||| |  | |   ||  |   || | ||| | | || || ||| ||| ||
Db    946 TTGGTTTTAGTGTTAATGTTGATACTCCCTCACTACTGGAAATATTTGCGCAAGAAAGTTTGGG     1005

Qy    989 ATGGTGATGAACCTTGCAAGTAGAATTCTTGCAAACACGAATAGTA------------  1033
          |||  ||| | ||    |||  |  || |   |  | || |  || |
Db   1006 CCTCTGAAATTACTCCTTTAGCAACTACACCAGCTGAAGATGCAACACCAACCAAGCTGCAT   1065

Qy   1034 ---TCACAGATGTTCTGTAGGGACTGTTTCTTTGGATTTTATTAGTGGAATTA   1090
             || |||| |  | | | |  |  | |    | |||  ||| || | |  |
Db   1066 CCTTAACAGATGTTTCATGAATCTATAGTAGTAGTGTCTGAAGCTAAATAACACATTA    1125

Qy   1091 GTTTTAGCAACTATGG---TCCATCAACTGGTTATTATATTTCCCTTATAAGTTGGTTA  1147
          ||||   ||| | |||   ||  |||| || | |  |||||   ||| | | |||| ||
Db   1126 GCTTTCGTTACTTTGGAGCTGAAAAAACAGCTTACTTTATTTCCCTTATAAATTAGTTA   1185

Qy   1148 AAGCAGCTGATGCTAATAACGTTGGATTACATAACGGTTACAATACAAGAAATGTTCAAC 1207
          |||      |       ||| || | | | ||||  || | ||||| | | || || ||
Db   1186 AAAC------TAGTGATAATGTTGGTTTACAATATAAGTTAAATGGTGGTGATACTAAAAC 1239

Qy   1208 AAGTTGAGTTTGCCACTTCAACTAGTGCAATAATAATACTAGCTAATCCAACTCCAGCAG 1267
```

```
Qy    760 ACTACAGTTAATGGAAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAAT   819
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACTACAGTTAATGGAAATATTAATAATACGTTATCAACTATTAATGAACAAAGACTAAT   540

Qy    820 GCTGATGCATTATCCTAATAGTTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTT  879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GCTGATGCATTATCCTAATAGTTTTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTT   600

Qy    880 GTAGGGACTTTTACAAACGCTAATGTTCAAACCTACAGTTTTGTTGCTTTTTAGT       939
          ||||||||||||||||||||||||||||||||||    ||||||||||||||||||
Db    601 GTAGGGACTTTTACAAACGCTAATGTTCAAACCTCAAACTACAGTTTTGTTGCTTTTTAGT  660

Qy    940 GCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTG              986
          ||||||||||||||||||||||||||||  |||||||| ||  ||||
Db    661 GCTGATGTAACACCCGTCAATTATAAAATGCAAGAGAACGGTTTG                 707

RESULT 15
AAT04075
ID   AAT04075 standard; DNA; 2196 BP.
XX
AC   AAT04075;
XX
DT   19-JUL-1996  (first entry)
XX
DE   M.gallisepticum DNA sequence I encodes 661 amino acid protein.
XX
KW   Detection; probe; primer; PCR; amplification; secretion; lung;
KW   avian chronic respiratory disease; respiratory tract; nasal cavity; ds.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key             Location/Qualifiers
FT   CDS             171..2156
FT   misc_difference 828..830
FT                   /*tag= a
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 978..980
FT                   /*tag= b
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 1083..1085
FT                   /*tag= c
FT                   /codon= seq: TGA, aa: Trp
FT   misc_difference 1911..1913
FT                   /*tag= d
FT                   /codon= seq: TGA, aa: Trp
FT                   /*tag= e
FT                   /codon= seq: TGA, aa: Trp
XX
PN   JP07236498-A.
XX
PD   12-SEP-1995.
XX
PF   25-FEB-1994; 94JP-00052764.
XX
PR   25-FEB-1994; 94JP-00052764.
XX
PA   (JAPG ) NIPPON ZEON KK..
PA   (SHIO ) SHIONOGI & CO LTD.
XX
DR   WPI; 1995-347462/45.
DR   P-PSDB; AAR79910.
XX
PT   Detection of Mycoplasma gallisepticum - for the quick detection, i.e.
PT   within one day, of avian chronic respiratory diseases.
XX
PS   Claim 7-10; Page 7-10; 11pp; Japanese.
XX
CC   This is the nucleotide sequence of a fragment of the Mycoplasma
CC   gallisepticum genome which codes for a 661 amino acid protein. This
CC   sequence and the sequence of AAT04076 (encoding a 261 amino acid protein)
CC   can be used to detect M.gallisepticum using probes based on nucleotides
CC   1125-1648 and primers based on nucleotides 449-466, the complement of
CC   bases 893-919, 1908-1934 and the complement of bases 2184-2210 of this
```

```
Db     1240 |||||| |||||   ||  ||  |  ||   |  ||    |  | | |  |  | | || ||
            AAATTAACTTTGTACAAACTCCAGCTTCTGGTTCAAGTGATGATGTTGCTGCTAATGAAGAAG 1299

Qy     1268 TTGATGAGATTAAAGTTGCTAAAATCGTTTTATCA 1302
               |   |   |   |||| ||| |   ||    ||
Db     1300 AAACTATGGCTAGTCCTGCTGAAATGCAGTCAGCA 1334
```

Search completed: May 5, 2004, 12:25:44
Job time : 463.318 secs

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     May 5, 2004, 10:52:31 ; Search time 98.4948 Seconds
                                     (without alignments)
                                     7358.420 Million cell updates/sec

Title:      US-09-901-572A-1
Perfect score:  1306
Sequence:   1 aaaaacatcagattgttaat.........taaaatcgtttatcaggtt 1306

Scoring table:  IDENTITY.NUC
            Gapop 10.0 , Gapext 1.0

Searched:   682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Issued_Patents NA:*
            1: /cgn2_6/ptcdata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptcdata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptcdata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptcdata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptcdata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptcdata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1280.4	98.0	2144	2	US-08-525-742-9	Sequence 9, Appli
2	1270	97.2	1387	2	US-08-525-742-3	Sequence 3, Appli
3	1264	96.8	1305	1	US-08-185-851A-1	Sequence 1, Appli
4	1086.8	83.2	2014	2	US-08-525-742-7	Sequence 7, Appli
5	216.6	16.6	2359	2	US-08-525-742-1	Sequence 1, Appli
6	66.4	5.1	6669	4	US-10-204-708-6	Sequence 6, Appli
c 7	65.6	5.0	3057	4	US-09-601-198-55	Sequence 55, Appli
8	61.2	4.7	5152	4	US-10-204-708-73	Sequence 73, Appl
c 9	59.4	4.5	11049	4	US-10-204-708-23	Sequence 23, Appl
10	55	4.2	10640	4	US-09-417-485D-5	Sequence 5, Appl
11	54	4.1	640681	4	US-09-790-988-1	Sequence 1, Appl
c 12	53.8	4.1	5476	4	US-10-204-708-82	Sequence 82, Appl
c 13	53.6	4.1	1887	4	US-09-601-198-39	Sequence 39, Appl
c 14	53.2	4.1	6113	4	US-10-204-708-14	Sequence 14, Appl
15	53.2	4.1	6317	4	US-10-204-708-12	Sequence 12, Appl
16	53.2	4.1	8700	2	US-08-645-193B-18	Sequence 18, Appl
17	52.8	4.0	7304	4	US-10-204-708-44	Sequence 44, Appl
18	52.4	4.0	429	4	US-09-543-681A-707	Sequence 707, App
19	52.4	4.0	1575	4	US-09-543-681A-615	Sequence 615, App
20	51.6	4.0	8700	2	US-08-392-625-16	Sequence 16, Appl
21	51.6	4.0	8700	2	US-08-466-961A-16	Sequence 16, Appl
22	51.6	4.0	53332	4	US-09-801-861-3	Sequence 3, Appli
c 23	51.2	3.9	6156	4	US-10-204-708-59	Sequence 59, Appl
24	51.2	3.9	6866	4	US-10-204-708-20	Sequence 20, Appl
c 25	50.8	3.9	1851	4	US-09-601-198-51	Sequence 51, Appl
c 26	50.6	3.9	14066	4	US-09-601-198-56	Sequence 56, Appl
c 27	50.2	3.8	7664	4	US-10-204-708-83	Sequence 83, Appl
28	50	3.8	2043	4	US-09-601-198-181	Sequence 181, App
c 29	49.8	3.8	8093	4	US-10-204-708-31	Sequence 31, Appl
c 30	49.4	3.8	615	3	US-08-998-416-186	Sequence 186, App
c 31	49.4	3.8	6866	4	US-10-204-708-20	Sequence 20, Appl
c 32	49.4	3.8	11049	4	US-10-204-708-21	Sequence 21, Appl
c 33	49.2	3.8	837	3	US-08-998-416-288	Sequence 288, App
34	49.2	3.8	8920	2	US-08-446-855A-1	Sequence 1, Appli
35	49.2	3.8	8920	3	US-09-150-741-1	Sequence 1, Appli
36	49.2	3.8	640681	4	US-09-790-988-1	Sequence 80, Appli
37	49	3.8	8961	4	US-10-204-708-80	Sequence 1137, Ap
38	48.6	3.7	636	3	US-08-998-416-1137	Sequence 72, Appl
c 39	48.6	3.7	8607	4	US-10-204-708-72	Sequence 43, Appl
c 40	48.4	3.7	7304	4	US-10-204-708-43	Sequence 36, Appl
c 41	48.2	3.7	1500	4	US-09-601-198-36	Sequence 13, Appl
42	48.2	3.7	19124	4	US-08-487-826B-13	Sequence 11, Appl
c 43	48	3.7	6317	4	US-10-204-708-11	Sequence 13, Appl
c 44	48	3.7	19124	4	US-08-487-826B-13	Sequence 10, Appl
c 45	47.8	3.7	6070	4	US-10-204-708-10	

```
                            ALIGNMENTS

RESULT 1
US-08-525-742-9
; Sequence 9, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
;  APPLICANT: Saito, Shuji
;  APPLICANT: Ohkawa, Setsuko
;  APPLICANT: Saeki, Sakiko
;  APPLICANT: Ohsawa, Ikuroh
;  APPLICANT: Funato, Hirono
;  APPLICANT: Iritani, Yoshikazu
;  APPLICANT: Aoyama, Shigemi
;  APPLICANT: Takahashi, Kiyoohito
;  TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
;  TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;  TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;  TITLE OF INVENTION: AS USE THEREOF
;  NUMBER OF SEQUENCES: 51
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;   ADDRESSEE: NAUGHTON
;   STREET: 1725 K Street, Suite 1000
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20006
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/525,742
;   FILING DATE: 25-SEP-1995
;   CLASSIFICATION: 424
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-074139
;   FILING DATE: 31-MAR-1993
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-245625
;   FILING DATE: 30-SEP-1993
;  PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP94/00541
;   FILING DATE: 31-MAR-1994
;  ATTORNEY/AGENT INFORMATION:
;   NAME: McLeland, Le-Nhung
;   REGISTRATION NUMBER: 31,541
;   REFERENCE/DOCKET NUMBER: 950811
;   TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 202-659-2930
```

```
;     TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO:    9:
;   SEQUENCE CHARACTERISTICS:
;     TYPE:  nucleic acid
;     LENGTH:  2144 base pairs
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  202..2046
US-08-525-742-9

Query Match        98.0%;   Score 1280.4;  DB 2;  Length 2144;
Best Local Similarity  98.8%;  Pred. No. 2.2e-262;
Matches 1290;  Conservative   0;  Mismatches   16;  Indels   0;  Gaps   0;

Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAACACAAAAATCTTCTAACAA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAACACAAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAACTAAAAAAATTAAAAAAAGTGTTTTTTCTTATC 120
            ||||||||||||||| |||||||||||||||||||||| |||| ||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAACTAAAAAAATTAAAAAAAGTGTTTTTTCTTATC 120

Qy    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTTATTTTTAGTCATCTTTTAAGAT 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTTATTTTTAGTCATCTTTTAAGAT 180

Qy    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAAGAAATCATCTTAAAGACTATTAGTTTG 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAAGAAATCATCTTAAAGACTATTAGTTTG 240

Qy    241 TTAGTACAACATCCTTTCTTAGCATTGGAATTCTAGTGTATGTCTATTACTAAAAAA 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTAGTACAACATCCTTTCTTAGCATTGGAATTCTAGTGTATGTCTATTACTAAAAAA 300

Qy    301 GATGCAAACCCAAATAATGCCAAATCCAATTAGAAGCAGCGCGAATGGAGTTAACAGAT 360
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    301 GACGCAAACCCAAATAATGCCAAATCCCAATTACAAGCAGCGCGCGAATGGAGTTAACTGAT 360

Qy    361 CTAATCAATGCTAAAGCCATGGATGACATTAGCTCACTACAAGACTATGCCAAGATTGAAGCT 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTAATCAATGCTAAAGCAAGGACAATTAGCTCACTACAAGACTATGCCAAGACTATGCCAAGATTGAAGCT 420

Qy    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAAATGCAACATTA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAAATGCAACACTA 480

Qy    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy    541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAACC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGCACTAAAAAACC 600

Qy    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATATCAAATT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTAGCTTCAACTGCTTATATCAGATT 660

Qy    661 CGCAATAATTTAGTGGATCTATACATATAATATGCTAGTAGTTTAATAACTAAAAACACTAGAT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGTAATAATTTAGTGGATCTATACATATAATATGCTAGTAGTTTAATAACTAAAAACACTAGAT 720

Qy    721 CCACTAAATGGGGGAACGCTTTAGATTCAATGAGATTACTACAGTTAATCGGAATATT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCACTAAATGGGGGAATGCTTTAGATTCAATGAGATTACTACAGTTAATCGGAATATT 780

Qy    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTTAATAGT 840
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAARGCTGATGCATTATCTTAATAGT 840
```

```
Qy    841 TTTATTAAAAAGTGATTCAAATAATGACAAAGTTTTGTAGGGACTTTTACAAACGCT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TTTATTAAAAAGTGATTCAAATAATGAACAAAGTTTTGTAGGGACTTTTACAAACGCT 900

Qy    901 AATGTTCAACCTCAAACTACACAGTTTTTGTCTTTTAGTGTGTGATGTAACACCCGTCAAT 960
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AATGTTCAACCTCAAACTACACAGTTTTTGTCTTTTAGTGTGTGATGTAACACCCGTCAAT 960

Qy    961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTGCA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTGCA 1020

Qy   1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGTCGGAACAAACACGAAG 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTTAGTCGGAACGAACACGAAG 1080

Qy   1081 TACCAATTAGTTTAGCAACTAGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACCAATTAGTTTAGCAACTAGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140

Qy   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATCAAATTAAATAATGGAAAT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATATATACTACAGCTAATCCAACT 1260
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATATATACTACAGCTAATCCAACT 1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
            ||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CCAGCAGTTGATGAGATTAAAATCGTTTTATCAGGTT 1306
```

```
RESULT 2
US-08-525-742-3
; Sequence 3, Application US/0825742
; Patent No. 5871742
; GENERAL INFORMATION:
;   APPLICANT:  Saito, Shuji
;   APPLICANT:  Ohkawa, Setsuko
;   APPLICANT:  Saeki, Sakiko
;   APPLICANT:  Ohsawa, Ikuroh
;   APPLICANT:  Funato, Hirono
;   APPLICANT:  Iritani, Yoshikazu
;   APPLICANT:  Aoyama, Shigemi
;   APPLICANT:  Takahashi, Kiyoohito
;   TITLE OF INVENTION:  NEW POLYPEPTIDE, DNA ENCODING THE
;   TITLE OF INVENTION:  POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;   TITLE OF INVENTION:  RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;   TITLE OF INVENTION:  AS USE THEREOF
;   NUMBER OF SEQUENCES:  51
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;     ADDRESSEE:  NAUGHTON
;     STREET:  1725 K Street, Suite 1000
;     CITY:  Washington
;     STATE:  DC
;     COUNTRY:  USA
;     ZIP:  20006
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0,  Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/525,742
;     FILING DATE:  25-SEP-1995
;     CLASSIFICATION:  424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  JP 05-074139
;     FILING DATE:  31-MAR-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  JP 05-245625
```

```
;       FILING DATE:  30-SEP-1993
;  PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/JP94/00541
;       FILING DATE:  31-MAR-1994
;  ATTORNEY/AGENT INFORMATION:
;       NAME:  McLeland, Le-Nhung
;       REGISTRATION NUMBER:  31,541
;       REFERENCE/DOCKET NUMBER:  950811
;  TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-659-2930
;       TELEFAX:  202-8870357
;  INFORMATION FOR SEQ ID NO:  3:
;  SEQUENCE CHARACTERISTICS:
;       LENGTH:  1387 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;  MOLECULE TYPE:  cDNA
;  FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  202..1305
US-08-525-742-3

Query Match        97.2%;  Score 1270;  DB 2;  Length 1387;
Best Local Similarity  98.8%;  Pred. No. 3.2e-260;
Matches 1290;  Conservative  0;  Mismatches 15;  Indels 1;  Gaps 1;

Qy    1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy   61 AATCCTAAATAAATAAGCCGTTAAATAACTAAGAAAAATTAAATTTTTATTTTTCTTATC 120
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAGAAAAATTAAATTTTTATTTTTCTTATC 120

Qy  121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTAAGAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTAGTCATCTTTAAGAT 180

Qy  181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAGACTATTAGTTTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAGACTATTAGTTTG 240

Qy  241 TTAGGTACAACATCCTTTCTTAGCATTGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TTAGGTACAACATCCTTTCTTAGCATTGGATTTCTAGCTGTATGTCTATTACTAAAAAA 300

Qy  301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGAATGGAGTTAACAGAT 360
          || ||||||||||||||||||||||||||||||  |||||||||||||||||||| |||
Db  301 GACGCAAACCCAAATAATGGCCAAACCCAATTACAAGCAGCCGAATGGAGTTAACTGAT 360

Qy  361 CTAATCAATGCTAAAGGCGGATGACATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCT 420
          ||||||||||||||||||  ||||||||||||||||||||||||||| ||||||||||||
Db  361 CTAATCAATGCTAAAGGCAAGGACATTAGCTTCACTACAAGACTATGCTAAGATTGAAGCT 420

Qy  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTGAAACAGTTAACAATAACCTTAACTGCAACATTA 480
          |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||| |||
Db  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAACTGCAACACTA 480

Qy  481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy  541 AAAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db  541 AAAAACGACTTTTGATAATGAACATCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600

Qy  601 ACTTTAGAAACAACGTGCTACTAACCTTGAAGGTTAGCTTCAACTGCTTATAATCAAATT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  601 ACTTTAGAAACAACGTGCTACTAACCTTGAAGGTTAGCTTCAACTGCTTATAATCAGATT 660

Qy  661 CGCAATAATTTAGTGGATCTATACAATAAATGCTAGTAGTTAATAACTAAAACACTAGAT
          ||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  661 CGTAATAATTAGTGGATCTATACAATAAATGCTAGTAGTTAATAACTAAAACACTAGAT 720
          ||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  721 CCACTAAATGGGGAACGCTTTAGATTCTAATGACTACACAGTTAATCGGAATATT 780
          |||||||||||||||  ||||||||||||||||| || |||||||||||||||||||
Db  721 CCACTAAATGGGGAATGCTTTAGATTCTAATGAGATTACACAGTTAATCGGAATATT 780

Qy  781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGTCATGTCATTATCTAATAGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  781 AATAATACGTTATCAACTATTAATGAACAAAAGACTAATGTCATGTCATTATCTAATAGT 840

Qy  841 TTTATTAAAAAGTGATTCAAATATAATGAACAAAGTTTGTGAGGGACTTTACAAACGCT 900
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  841 TTTATTAAAAAGTGATTCAAATAATAATGAACAAAGTTTGTGAGGGACTTTACAAACGCT 900

Qy  901 AATGTTCAACCTTCAAACTACACAGTTTGTTGCTTTGTGCTTTAGTGCTGATGTAACACCCGTCAAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AATGTTCAACCTTCAAACTACACAGTTTGTTGCTTTGTGCTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy  961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAGTAGAATTCTTGCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGAACCTTCAGTAGAATTCTTGCA 1020

Qy 1021 AACACGAATAGTATCACAGATGTTTCTTGGATTATAGTTAGCTGGAACAAACACGAAG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 AACACGAATAGTATCACAGATGTTTCTTGGATTATAGTTAGCTGGAACAAACACGAAG 1080

Qy 1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||| ||
Db 1081 TACCAATTTAGTTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAG 1140

Qy 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGAAAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGAAAT 1200

Qy 1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAATAATACTACAGCTAATCCAACT 1260

Qy 1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
          |||||||||||||||||||||||||||||||||||||||||||||
Db 1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1305
```

RESULT 3
US-08-185-851A-1
; Sequence 1, Application US/08185851A
; Patent No. 5489430
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Fujisawa, Ayumi
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
; TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
; TITLE OF INVENTION: Well As Vaccines Utilizing the Same
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K Street, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
; SOFTWARE: ASCII from Word Perfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,851A
; FILING DATE:

```
;     CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;     NAME: Theresa M. Stevens-Smith
;     REGISTRATION NUMBER: 36,281
;     REFERENCE/DOCKET NUMBER: PO-8-A930918
; TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 202-659-2930
;     TELEFAX: 202-887-0357
; INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 1305 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: circular
; HYPOTHETICAL: NO
;     FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 202..1308
US-08-185-851A-1

Query Match     96.8%;  Score 1264;  DB 1;  Length 1305;
Best Local Similarity  98.3%;  Pred. No. 5.9e-259;
Matches 1284;  Conservative  0;  Mismatches  21;  Indels  1;  Gaps  1;

Qy    1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy   61 AATCCTAAATAATAAGCCGTTAAATTAACTAAAAAAAGTTAAAAAAAATGGTTTTTCTTATC 120
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AATCCTAAATAATAAGCCGTTAAATTAACTAAAAAAAGTTAAAAAAAATGGTTTTTCTTATC 120

Qy  121 AACCAAAATTCTCTAGTAATAAACGCTATTATTTTTATTTTTAGTCATCTTTAAGAT 180
          |||||||||||||||||||||||||| |||||||||||||| |||||||||||||||
Db  121 AACCAAAATTCTCTAGTAATAAACGCTATTATTTTTATTTTTAGTCATCTTTAAGAT 180

Qy  181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy  241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAA 300

Qy  301 GATGCAAACCCAAATAATGCCAAACCCAATTAGAAGCAGCGCGAAATGGAGTAAACAGAT 360
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||| |||
Db  301 GACGCAAACCCAAATAATGCCAAACCCAATTACAAGCAGCGCGAAATGGAGTAACTGAT 360

Qy  361 CTAATCAATGCTAAAGCGATGAAGCTCACTCACAGACTATGCCAAGATTGAAGCT 420
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CTAATCAATCTAAAGCAAGGACACATTAGCTCACTCACAGACTTAAGATTGAAGCT 420

Qy  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAGCTGAAACAGTAACAATAACCTTAATGCAACATTA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AGTTTATCATCTGCTTATAGTGAAGCTGAAGCTGAAACAGTAACAATAACCTTAATGCAACACTA 480

Qy  481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCGCCATCAACAAGCTAATACGGAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 GAACAACTAAAAATGGCTAAAACTAATTAGAATCAGCGCCATCAACAAGCTAATACGGAT 540

Qy  541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 AAAACGACTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACC 600

Qy  601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATT 660
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  601 ACTTTAGAACAACGTGCTACTAACCTTCAACTGCTTAGCTTCAACTGCTTATAATCAGATT 660

Qy  661 CGCAATAATTTAGTGGATCGTCATCATACACATAAAGCTAGTAGTTTAATAACTAAACACTAGAT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 CGTAATAATTTAGTGGATCGTCATCATACACATAAATGCTAGTAGTTTAATAACTAAACACTAGAT 720

Qy  721 CCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTCTAATGAGATTACTACAGTTAATCGGAATATT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CCACTAAATGGGGGAATGCTTTTAGATTCTAATGAGACTAATGAGATTACTACAGTTAATCGGAATATT 780

Qy  781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATATCTAATAGT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCATTATCTAATAGT 840

Qy  841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 TTTATTAAAAAGTGATTCAAAATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCT 900

Qy  901 AATGTTCAACTTCAAACTACACAGTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 AATGTTCAACCTCAAACTACACAGTTTTGTTGCTTTTAGTGCTGATGTAACACCCGTCAAT 960

Qy  961 TATAAATATGCAAGAAGGACCGTTGGAATGGTGATGATGAACCTTCAGTAGAATTCTTGCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TATAAATATGCAAGAAGGACCGTTNNNAATGGTGATGATGAACCTTCAGTAGAATTCTTGCA 1020

Qy 1021 AACACGAATAGTATCACAGATGTTTCTTGGATTTATAGTTAGCTGGAACAAACAGAAG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 AACACGAATAGTATCACAGATGTTTCTNNNNATTTATAGTTAGCTGGAACAAACACGAAG 1080

Qy 1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTATATTTCCCTTATAAG 1140
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db 1081 TACCAATTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTTCCCTTATAAG 1140

Qy 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db 1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATTAAATAATGGAAAT 1200

Qy 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACCAGTGCAAATAATATTACACAGCTAATCCAACT 1260
          ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db 1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATATTACACTAGCTAATCCAACT 1260

Qy 1261 CCAGCAGTTGATGAGATTAAAAGTTGCTAAAATCGTTTTATCAGGTT 1306
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 -CAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTTATCAGGTT 1305
```

RESULT 4
US-08-525-742-7
```
; Sequence 7, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
;     APPLICANT: Saito, Shuji
;     APPLICANT: Ohkawa, Setsuko
;     APPLICANT: Saeki, Sakiko
;     APPLICANT: Ohsawa, Ikuroh
;     APPLICANT: Funato, Hirono
;     APPLICANT: Iritani, Yoshikazu
;     APPLICANT: Aoyama, Shigemi
;     APPLICANT: Takahashi, Kiyoohito
;     TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
;     TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
;     TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
;     TITLE OF INVENTION: AS USE THEREOF
;     NUMBER OF SEQUENCES: 51
;     CORRESPONDENCE ADDRESS:
;     ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
;     ADDRESSEE: NAUGHTON
;     STREET: 1725 K Street, Suite 1000
;     CITY: Washington
;     STATE: DC
;     COUNTRY: USA
;     ZIP: 20006
; COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/525,742
```

```
;   FILING DATE: 25-SEP-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-074139
;   FILING DATE: 31-MAR-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 05-245625
;   FILING DATE: 30-SEP-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP94/00541
;   FILING DATE: 31-MAR-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: McLeland, Le-Nhung
;   REGISTRATION NUMBER: 31,541
;   REFERENCE/DOCKET NUMBER: 950811
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-659-2930
;   TELEFAX: 202-8870357
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2014 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 54..1883

US-08-525-742-7

Query Match        83.2%;   Score 1086.8;   DB 2;   Length 2014;
Best Local Similarity 96.7%;   Pred. No. 2.1e-221;
Matches 1120;   Conservative   0;   Mismatches  37;   Indels   1;   Gaps   1;

Qy    150  TTTATTTTATTTTAGT-CATCTTTTAAGATAATAAATATATCTTAATAATATTCTATGAATA  208
           ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTATTTTATTTTGGTAAATCTTTTAAATATAAATAATAATATTTTAATAATTCTATGAATA   60

Qy    209  AGAAAAGAATCATCTTAAAGACTATTAGTTGTTGTTAGTACAACATCCTTCTTAGCATTG  268
           |||||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||
Db     61  AAAAAAGAATCATCTTAAAGACTATTAGCTTGTTGTTAGTACAACATCCTTCTTAGTATTG  120

Qy    269  GGATTCTCTAGCTGTATGTCTATTACTTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC  328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GGATTCTCTAGCTGTATGTCTATTACTTAAAAAAGATGCAAACCCAAATAATGGCCAAACCC  180

Qy    329  AATTAGAAGCAGCGGAATGAGAGTTAACAGATCTAATCAATGCTAAAGCTAAAGCTGACATTAG  388
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    181  AATTAGAAGCAGCGGAATGAGAGTTAACAGATCTAATCAATGCTAAAGCTAAAGCTGACATTAG  240

Qy    389  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  448
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  300

Qy    449  AAACAGTTAACAATAACCTAATGCAACAACTAGAACAACATTAGAAATGGCTAAAACTAATT  508
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AAACAGTTAACAATAACCTAATGCAACAACTAGAACAACATTAGAAATGGCTAAAACTAATT  360

Qy    509  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAA  568
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAA  420

Qy    569  ATTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTG  628
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  ATTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTG  480

Qy    629  AAGGTTTGTTGTCATCAACTGCTTATAATCAAATTCGCAATAAATTAGTTGGATCTATACAATA  688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  AAGGTTTGTTGTCATCAACTGCTTATAATCAAATTCGCAATAAATTAGTTGGATCTATACAATA  540

Qy    689  AAGCTAGTAGTTTAATAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCGCTTTAGATT  748
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  AAGCTAGTAGTTTAATAATAACTAAAACACTAGATCCACTAAAATGGGGAACGCGCTTTAGATT  600

Qy    749  CTAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAAC  808
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTAATGAGATTACTACAGTTAATAAGAATATTAATAATACGTTATCAACTATTAATGAAC  660

Qy    809  AAAAGACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATG  868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AAAAGACTAATGCTGATGCATTAGCTAATAGTTTTATTAAAGAACTGATTCAAAATAATA  720

Qy    869  AACAAAGTTTGTAGGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTG  928
           |||||||||||||||||||||| ||||||||||||||||| ||||||||||||||||
Db    721  AACAAAGTTTGTAGGAATGTTACAAACTAATGTTCAACCTTCAAACTATAGTTTTTG  780

Qy    929  TTGCTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGA  988
           ||||||||||||||||||||||||||||||| || ||||||||||||||||||||||||
Db    781  TTGCTTTAGTGCTGATGTAACACCTGTTAATTATTATAAATATGCAAGAAGACGGTTTGAA  840

Qy    989  ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTT  1048
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  ATGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACCAATAGTATTACTGATGTTTCAT  900

Qy   1049  GGATTTATAGTTAGCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATGGTC  1108
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    901  GAATTTATAGTTTATCTGGAACAAACAAACGAAATACCAATTTAGTTTTAGCAACTACGGTC  960

Qy   1109  CATCAACTGGTTATTTATTTTCCCTTATAAGTTGGTTAAAGCAGCTAATGCTAATAACG  1168
           ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    961  CATCAACTGGTTATTTATTTTCCCTTATAAGTTTAAAGCGGCTCGATGCTGATGTAATAATG  1020

Qy   1169  TTGGATTACAATACAATAATCAAATTAAATATGGAAAATGTTCAACAAGTTGCCACTTCAA  1228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  TTGGATTACAATACAATAATCAAATCAAATTAAATATGGAAAATGTTCAACCAGTTGCCACTTCAA  1080

Qy   1229  CTAGTGCAAATAATACTACAGCTAATCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA  1288
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  CTAGCCAAATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTA  1140

Qy   1289  AAATCGTTTTATCAGGTT  1306
           ||||||||||||||||||
Db   1141  AAATCGTTTTATCAGGTT  1158
```

RESULT 5
US-08-525-742-1
; Sequence 1, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyoohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

```
Db   889  ATTTGACTAGAGTTG---ATGTAGCAATATCAGGAGCAACCAGCAAATTACAGTTTTG  945
          |||| |||  | |||      ||  |||||  | ||| ||  |||  || || ||||||
Qy   929  TTGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGA  988

Db   946  TTGGTTTTAGTGTTGTTAATGTTGATACTCCTAACTGAAATTTTGCGCAAGAAAAGTTTGGG  1005

Qy   989  ATGGTGATGAACCTTCAAGTAGAGAATTCTTGCAAACACGAATAGTA-----------  1033
          ||  ||  |  ||   |||   |  ||| ||||  | ||  ||
Db  1006  CCTCTGAAAATACTCCTTAGCAACTACACCAGCTGAAGATGCAACACACAAGCTGCAT  1065

Qy  1034  ---TCACAGATGTTTCTTGGATTTATAGTTTAGCTGGAACAAAACACGAAGTACCAATTTA  1090
             || || || | |||   ||| | ||| ||||  |  || || ||  || || ||
Db  1066  CCTTAACAGATCGTTCATGAATCATAGTTTAAATGTGCTGAAGCTAAATACACATTAA  1125

Qy  1091  GTTTTAGCAACTATGG---TCCATCAACTGGTTATTTATATTTCCCTTATAAGTTGGTTA  1147
             ||| || ||| |||    || |  ||| ||| || ||| ||||||||||||  | |||||
Db  1126  GCTTTCGTTACTTTGGAGCTGAAAAAAACAGCTTACTTATTATTTCCCTTATAAAATTAGTTA  1185

Qy  1148  AAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAAATAATGGAAATGTTCAAC  1207
          ||      || |  || || ||| |  || ||||| | || || ||| | | |||
Db  1186  AAAC-------TAGTGATAATGTGGTTTACAATATAGTTAAATGTGGTGTATACTAAAAC  1239

Qy  1208  AAGTTGAGTTTGCCACTTCAACTAGTGCAAATAAATAATCACAGCTAATCCAACTCCAGCAG  1267
          ||  |  | |||  ||| | ||| || |||||| | || | || ||| ||| || || |
Db  1240  AAATTAACTTTGTACAAACTCCAGCTTCTGGTTCAAGTGATGTTGCTGCTAATGAAGAAG  1299

Qy  1268  TTGATGAGATTAAAGTTGCTAAAATCGTTTTATCA  1302
          |  | || | ||  |||||  |||| ||| | |||
Db  1300  AAACTATGGCCAGTCCTGCTGAAATGCCAGTCCAGCA  1334
```

```
Query Match          5.1%;  Score 66.4;  DB 4;  Length 6669;
Best Local Similarity 44.2%;  Pred. No. 2.7e-05;
Matches 363;  Conservative 0;  Mismatches 456;  Indels 3;  Gaps 2;

Qy   55  TAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATATTAAAAAATGGTTTTT  114
         ||   | || ||| ||||| |    ||| ||| | |||  |  | |||||||||    ||
Db 4745  TATCTACACATAAACAAATATTAAAAAATTAAAATCAAATTAAAAAATTAAAAAT  4686

Qy  115  CTTATCAACCAAAATTCTCTAGTAATAAACGTTTATTTATTTATTTTAGTCATCTTT  174
```

```
Query Match         16.6%;  Score 216.6;  DB 2;  Length 2369;
Best Local Similarity 55.1%;  Pred. No. 3.5e-37;
Matches 548;  Conservative 0;  Mismatches 414;  Indels 33;  Gaps 5;

Qy  329  AATTAGAAGCAGCGCGAATGAGCGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAG  388
         |||||| || | | | ||| | || ||||| |  |  || ||||  | |   | || | |||
Db  352  AATTAGCACACGCAAAAGCTGCTTAACTACTTTGATTAATGGTGAAACTGCAAATCTTG  411

Qy  389  CTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTG  448
         | || | || ||||||||| | || |||   || |||| |   ||| ||||| ||  | |
Db  412  CGTCATATGAAGACTATGCTAGATCAAAGTGAATTAACATCAGCGTATGAAAACAGCTA  471

Qy  449  AAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATT  508
         || || || ||  ||  |  |||||| | || |  || ||||  |||   || | | ||
Db  472  AAGCAGTTTCAGTTAAAACTGGTGCAACTCTAAATGAGGTTAATGAGGCAAAAACTACAT  531

Qy  509  TAGAATCAGCCATCAACCAAGCTAATACGGATAAAAGCTAAAGAACCACCCAA  568
         ||||  | || ||  |   |||   ||  | ||  | |||| | || ||| ||
Db  532  TAGATGCTGCTATTAAAAAAGCTGCTAGTGTGCTAAGTAAATGATTTTGATGCACAGCACGGGT  591

Qy  569  ATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCCTG  628
         | |||||| ||||||| ||| |||| |||| || |||  |||||||  |||| || | |
Db  592  CACTAGTGAAGCATATAACAATCTAAAACAACGTTAAAAGAAGAAAAACGTAAATTTAG  651

Qy  629  AAGGTTTGTCATCAACTGCCTTATAATCAAAATTCGCAATAATTTAGTGGATCTATACAATA  688
         |   | | ||||  | |  |  ||| ||||  | |||| |  | ||  ||| | |||||||
Db  652  ATTCTTCTTCAAACGAAAAATATGCAGCAATCAGAACAACTAATCTTAATAGTTTATATGAA  711

Qy  689  AAGCTAGTAGTTTAATAACTAAAAACACTAGAATCCACTAAGTGGGGAACGCTTTTAGATT  748
         ||||  || |||| |||| | |||||||   ||||| |||| |||| |||| ||| ||| |
Db  712  AAGCCAATACTATTGTTACAGCGCACTTTAGACCCTGCTACTGGAAATATTCCTGAAGTTA  771

Qy  749  CTAATGAGATTACTACAGTTAATCGGAATATTAATATAAATACGTTATCAACTATTAATGAAC  808
         | | || | || || | ||||||  |||  | || |  | |||||| || | ||| || || |
Db  772  TGAGTG---TAACACAAGCTAATCAAGATATTACTAATTACTAATGGTAAAATAATAGCTT  828

Qy  809  AAAAGACTAATGCTAATGCATTATCTAATAGTTTTATTAAAAAGAAGTAATCAAAATAATG  868
         ||||| | ||||||  | |  ||||  || |||| |   ||| ||   | || ||||| |
Db  829  GAAAACAAAATGCTGATAAATTAGCTAACAGTTTATCAAACAGTCTTTAGTTAAAAATA  888

Qy  869  AACAAAGTTTGTAGGGACTTTACAAACGCTAAATGTTCAACCTTCAAACTACCAGTTTTG  928
```

```
;  LENGTH: 3057
;  TYPE: DNA
;  ORGANISM: Ureaplasma urealyticum
US-09-601-198-55;

Query Match             5.0%;  Score 65.6;  DB 4;  Length 3057;
Best Local Similarity  46.2%;  Pred. No. 3.4e-05;
Matches 218;  Conservative  0;  Mismatches 254;  Indels  0;  Gaps  0;

Qy    444 AGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACACTAAAATGGCTAAAAC 503
                      |    ||| | || |||  |||||  |||||||   |  |||  ||||
Db   2350 ATCAGCTAAATTTAAAAATAAGCCAGATTTAGCTAAATGAATGTAACAATAAAATAAAAA 2291

Qy    504 TAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACA 563
                      || |  || | | |||||   | | | || |||  |    |||  |
Db   2290 TAATCTAATATATGGAGGTAATCTTAGATACTAAAATCAAAACATTAGTTGATAAAACTAAA 2231

Qy    564 CCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAA 623
                    ||||  |  | | ||| | | ||  | ||  ||| || | || || || |||
Db   2230 ATTATTGAAATTAGAATCAAATAATGATATTTATTACAATATACAAGAAAATGCTGTTAA 2171

Qy    624 CCTTGAAGGTTTGTCATCAACTGCTTATATACGCTTATATAATCAAATTCGCAATAATTTAGTGGGATCTATA 683
                  |  | |   ||  |||  || | |  ||| |||  |||   || |||| || |  ||||
Db   2170 TCCTAATAGAGTGATAGGAACAATAAAAGCAGAATTTTCAAGATTAATTCTAATTACAT 2111

Qy    684 CAATAAAGCTAGTAGTTTAATAACATAAAAACACTAGATCCACTAAAATGGGGGAACGCTTTT 743
                |    | | |  ||    |  || |    |  ||  | |   | | || ||| |  ||
Db   2110 TAACAAAAAATTAAATTAGTTATATGAAGATAACAATCATCAAGAATGAGAAACTAATGA 2051

Qy    744 AGATTCTAATGAGATTACTACAGTTAATCGGATTATTAATAATACGTTATCAACTATTAA 803
                 ||||  | |  | | | |||  ||||  | | |||| ||| || | || |||
Db   2050 ACTAACTTTATCAAACAATAAAACTAGTTATAATTTTGATATTAGTAATAAACTAACTCCTAA 1991

Qy    804 TGAACAAAGACTAATGCTGATGCATTATCTCATTAGTTTTATTAAAAAGTGATTCAAAA 863
                 ||  | |||||  | || || |   ||  |  || || ||| || || || | ||||
Db   1990 TAGAAAATATTCTTTCAAAAAGATTTAATTTCAAATAATAATATCTTTGAAGAATTGCT 1931

Qy    864 TAATGAACAAAGTTTGTAGGGACTTTTACCAACGCTAATGTTCCAACCTCA 915
                |||| | || ||| ||  | || ||| | |  || | |  | | |||| ||
Db   1930 AAATTCAAATAATAATTGCTGATTCATTTATATATCATGCACCAACTGCA 1879
```

RESULT 8
US-10-204-708-73/c
; Sequence 73, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
;   TITLE OF INVENTION: by Assessing DNA Methylation
;   FILE REFERENCE: 5013.1012
;   CURRENT APPLICATION NUMBER: US/10/204,708
;   CURRENT FILING DATE: 2003-05-06
;   PRIOR APPLICATION NUMBER: PCT/EP01/03971
;   PRIOR FILING DATE: 2001-04-06
;   PRIOR APPLICATION NUMBER: DE 10019058.8
;   PRIOR FILING DATE: 2000-04-06
;   PRIOR APPLICATION NUMBER: DE 10019173.8
;   PRIOR FILING DATE: 2000-04-07
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 73
;   LENGTH: 5152
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73

```
Db   4685 CAATTCCACAAAACACCTTAAATAAATAAATAAAATAAAAACTAAACTAATAACCAAAAACAAC 4626
                |||  |||  |  || |||  || |||  | |||  | |||||  ||||  |||| |||
Qy    175 TAAGATAAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATT 234
                ||| || ||  |||  | || | |    | || | ||  |||   |||||
Db   4625 AAACACTTTAATATATTCTAACCCACACAACATTTTATAAAATCCCAACAAACAATAAA 4566
                 ||||| ||  | ||| ||| ||| || ||||| |  ||| |  |||||| |||
Qy    235 AGTTTGTTAGGTACAGACATCCTTCTTAGCATGGATTTCTAGCTGTGTATGTCTATTACT 294
                |   || ||  |  || ||  |||  | || || |  || | ||| ||  ||||||
Db   4565 AACTACTTAAATAAATAATAAATACAAAAATCAAAATCCTAAAAATAAAA 4506
                 |||| || ||| |  | | || |||   | | |||  ||||||||
Qy    295 AAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAACGCAGCGCGGAATGGAGTTA 354
                |||| |  |  |  | ||  |||||  |  ||| | ||||
Db   4505 AATAAAAAACAACAACTAAAATCATTATTCTTTAAAATTAAACTTATA 4446
                 ||  || |||| ||| | | ||||| | |||  |||  | ||| |||||
Qy    355 ACAGATCTAATCAATCGCTAAAGCGATGACATTAGCTTCACTACACAAGACTATGCCAAGATT 414
                | ||  | || || |  |  |  | |  ||| | | ||||| |||
Db   4445 AAAAAATAACACAAAACATAAACAAAATATTCTATTTAATAAATAAATCTAATATATCTAATA 4386
                 |||||  ||||  ||||| ||||  | |||| |||  | | || ||||
Qy    415 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA 474
                 ||| |||  ||||| |  || ||  |||
Db   4385 AAATCCAAAAAAAATATTTTTTATATATAATATTTAAAAAAAAACCTAATTATC 4326
                ||  | ||| || || | ||  | || || |  |  ||||||||
Qy    475 ACATTAGAACAACTAAAAATGCTAAAAACTAATTTAGAATCAGCCATCAACCAAGCTAAT 534
                 | || || | ||  |  |||   ||||   | |||| | ||||
Db   4325 TAAACACTAAAAAAAAAACTTAACTTCTCTATATTCAACAATTAACCACAATATTTAA 4266
                 ||||  || | || |||   || | ||| | | |||| || ||||||||
Qy    535 ACGGATAAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 594
                |  | | || | || | | | | || | |  | ||  | || |||||
Db   4265 ACATACTAAACACCCAAAATTTACTAAAAAAAAACATTATAAAAAAACCAATAACAAA 4206
                 || | |||  | ||  |||| |||  | |||| || ||||
Qy    595 AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATAT 654
                 || |  ||||  ||  ||||||| | ||||   |  |  ||  || |  |||||
Db   4205 AAAAAACTAAAAAATAACGAAAAAAATACTAATAAAAAAAAATCATAAAAATAATAAAACC 4146
                 ||||  ||| | |||| ||| | ||||| || |  |||
Qy    655 CAAATTCGCAATAATTTAGTGGGATCTATA--CAATAAAGCTAGTAGTTTAATAACTAAAA 712
                ||  || | |||| |   |   ||| |||   | || ||| || || |||  ||||||
Db   4145 TCATATAAAAAAAATATACACGATATATATCTATTTAAAATAAATAATAATAAAACGACAA 4086
                ||| || |||||| |  || ||| || ||  |||| |||||
Qy    713 CACTAGATCCACTAAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTAATC 772
                 ||  ||| | | |||  || || ||||| | |||
Db   4085 AAATAAAAACATTAAAACATAAAAAAAACTTTTTTATATTAATTACTCCTACCCATCTAAAA 4026
```

```
Qy    773 GGAATTAATAATACGTTATCAACTATTAAGTGAACAAAGACTAATGCTGATGCATTAT 832
                 |||  ||||| |||| | ||| |  | | || |  || || |
Db   4025 TTTTTAATAAAAATAAAATTATCTCAACTTTCAACTTTCAAACTAAAAAAT-ATAATATAAAAACTT 3967
                |||  |||||  |||| ||| ||| |||||| ||| || |||  ||||||||
Qy    833 CTAATAGTTTTATTAAAAAAGTGATTCAAAAATAAATGAACAAA 874
                ||||||  | |||| | |  |||| |  |||| |  || ||
Db   3966 CAAAAAAATTTATAACAACAATAAAATAAAATTACGAAA 3925
```

RESULT 7
US-09-601-198-55/c
; Sequence 55, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;   APPLICANT: Cassell, Gail H.
;   APPLICANT: Chen, Ellson Y.
;   APPLICANT: Glass, Jennifer S.
;   APPLICANT: Glass, John I.
;   APPLICANT: Heiner, Cheryl R.
;   APPLICANT: Lefkowitz, Elliot
;   TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;   TITLE OF INVENTION: UREALYTICUM
;   FILE REFERENCE: UAB-13452/22
;   CURRENT APPLICATION NUMBER: US/09/601,198
;   CURRENT FILING DATE: 2000-12-08
;   PRIOR APPLICATION NUMBER: 60/073,189
;   PRIOR FILING DATE: 1998-01-30
;   NUMBER OF SEQ ID NOS: 181
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55

Query Match
Best Local Similarity   4.7%;   Score 61.2;   DB 4;   Length 5152;
              41.9%;   Pred. No. 0.00033;
Matches 429;   Conservative 0;   Mismatches 593;   Indels 2;   Gaps 1;

```
Qy   24 ATATCTTTGCTTAAAAAACACAAAAATCTTCTACACAAAATCCTAAATAAATAAGCCGTTA  83
         ||||||||||                              ||||||||||||||||||
Db 4550 ATAATACTTCTAAAAATCCTAAATAAATAAATAATTATCCTAAATACTAAAAACTAC    4491

Qy   84 AATTACTAAAAAAATTAAAAAAATGGTTTTTCTTATCACCAAAATCTCTAGTAATAAA    143
         ||||                    |||||||||||||||||||
Db 4490 CACATATTACTTTTATAAACCTTTTTTTTTTTTTAACTATTTTTATTAAAAATAAA    4431

Qy  144 CGCTTATTTATTTTTATTTTTAGTCATCTTTAAGATATAAATATATCTTAATATTCTAT    203
         ||          |||||||                 ||||||
Db 4430 CCTAAAATAAAATTCAAAACACTTTTTAAATTAAAATAAAACCTATTTCTAAATACCA-   4372

Qy  204 GAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACCAACATCCTTTCTTAG   263
         ||||     ||||                      ||||
Db 4371 -AATTATTATTTATTTTAAACAATAAATATTTTACAACAAAACATATATACCGAA        4313

Qy  264 CATTGGGATTTCTAGCTGTATGTCTATTACTATCTAAAAAAGATGCAAACCCAAATAATGGCCA  323
         ||||          |||||||||||          ||||||||
Db 4312 CATAATAACTCATAACCTATAATCCCAACACTTTAAAAAAACCGAAATAAATAAAACACTTA    4253

Qy  324 AACCCAATTAGAAGCAGCGCGAATGGAGTTAACGACATCTAATCAATGCTAAGGCCGATGAC   383
         ||          |||||          ||||||||
Db 4252 AATCCAAAAATTCCAAACCATCCTAAACAATAATAAAAACCCTATCTTCTTACAAAATAAA    4193

Qy  384 ATTAGCTTCACTCACAGACTATGCCAAGATTGAAGCTAGTTTATCATCTCGCTTATAGTGA   443
         ||                                               ||||
Db 4192 AAAAAAAAATAATAAAAATAAAATTAACCAAATAAAAATCACTTACCTACA            4133

Qy  444 AGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAATGGCTAAAAC     503
         ||||                    ||||||
Db 4132 ATCCCAACTACTCAAAATCTCAAACAAAAAATCTAAACCCTATCTCA               4073

Qy  504 TAATTTAGTTGAATCAGCCATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACA  563
         ||||          ||||||||
Db 4072 TCATAAACCGTAATTACGCCCACTACATTTCAAACCTAAACAAAAACCCTATCTCTA       4013

Qy  564 CCCAAATTTAGTTGAAGCATACAAAGACTAAAGAGACAACGTGCTACTAA            623
         ||||                             ||||
Db 4012 AAAAAAAAAAAAAAAAAAAAATTTAAGAACAATAAAGGCCGCCATCATATAA          3953

Qy  624 CCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTTGGATCTATA   683
         ||||          |||||||||
Db 3952 CTATTATTATACTTATATTAAAAAAATACTTTAAAATTTTTTTATTTTTATTTTTATTT    3893

Qy  684 CAATAAAGCTAGTAGTTTAATAATACTATAAAACACTAGATCCACTAAAATGGGGAACGCTTTT 743
         ||                                    ||||
Db 3892 ATCTTAAATACTTTACTTACTTACTTATTTAACTATTTCAAAATTATTACAAAGAACAT     3833

Qy  744 AGATTCTAATGAGATTACTACAGTTAATCGGAATAATAATAAAACGTTAATCAACATTAA    803
         ||||                             ||||
Db 3832 AAAATCAATTTTCTTTAACACCTAAATTAAAAAATAAAAAAATATTTCATTAAAACTAA     3773

Qy  804 TGAACAAAAGACTAATGCCTGATGATGCATTATCATTAATAGTTTTATTAAAAAGTGATTCAAAA 863
         ||                        |||||||
Db 3772 TTTCCTTTAAACACAATCTTAATTAATATAAAAAATTTTCCCATTTTTAAAAAAATTCCTAA   3713

Qy  864 TAATGAACAAAGTTTGTAGGGACTTTTACAAACGCTAATGTTCAAACCTTCAAACTTACACAG  923
         ||                                    |||||||||
Db 3712 AATACCTAAATATTAAAAACACTACTTACACTTCTACACTTCTTACAAACTATAT         3653

Qy  924 TTTTGTTGCTTTTAGTGCTGATGTGAACCTTCCTGCAAACACGAATTCTTGCAAACAGGACCGT 983
         ||||||
Db 3652 TTTACTTATTTTCTTTACATTTACAAAAATCAAAAATTTTTAAAACGATTATTTCCAACTC    3593

Qy  984 TTGGGATGGTGATGAACCTTCAGTGAGTGATGAACCTTCAGTGAACAAACGAATAGTATCACAGATGT 1043
         ||||                                    ||||
Db 3592 TTATTTTAAAATTTAAATATAATAAAATTAAACTAAAAAAAAACACTTTATT            3533
```

```
Qy 1044 TTCT 1047
         ||||
Db 3532 TTCT 3529
```

Query Match   4.5%;   Score 59.4;   DB 4;   Length 11049;
Best Local Similarity 43.6%;   Pred. No. 0.00092;
Matches 267;   Conservative 0;   Mismatches 346;   Indels 0;   Gaps 0;

```
Qy    1 AAAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAACACAAAATCTTCTAACAA    60
         ||||||||||                                          |||||||
Db 1317 AATAACTTAAAATTAAAATAATTAAAAAAACAAAATAAATATTAAAATAAAACGTTAACTC 1258

Qy   61 AATCCTAAATAAATAAGCCGTTAAATTAACAATAAAAATTAAAAAAATGGTTTTCTTATC   120
         ||||          ||||                            ||||
Db 1257 AAACCTCCTTTCACTATCATAAATTTTCTAACTAATATAATAATTTCATT            1198

Qy  121 AACCAAAATTCTCAGTAATAAACGCTTATTTTATTTTTAGTTCATCTTTAAGAT        180
         ||                                                  ||||
Db 1197 ATCTTAAATACTTTACTAATTAACTATTATTTAACTACTATAACGATTATTCATCCTTACGCCT 1138

Qy  181 ATAAATATATCTTAATATTCTATGAATATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG   240
         ||                                          ||||
Db 1137 TTATTTCACATTCGTATTCATTCAATATAAAAATTCTCCCAATAACCTACAAAAACAAAAATC 1078

Qy  241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTGTATGTCTATTACTATAA   300
         ||||                                          ||||
Db 1077 TAAATTAACTCTCACTTTCAAAACTTTCAAAATAAAAAAATCATAAATAAATCATAAAACCGATTTAA 1018

Qy  301 GATGCAAACCCAAATAATGGCCAAATAATGGCCAAATTAGAAGCAGCGCGGAATGAGTTAACAGAT 360
         ||||                                          ||||
Db 1017 TTCTCTAATCCTTATAAAATTCTAATATTCTAATTCTTACAAGACTATGCCAAGATTGAAGCT  958

Qy  361 CTAATCAATGCTAAAGCCTAAAGCCGATGACATTAGCTTTATCATCATGCCAAGAATTGAAGCT 420
         ||||                                          ||||
Db  957 AATAAAATACTTTATAAAAATTATAAAAATCAAAAAACAATAAACCATATTTATTTAAAAACT   898

Qy  421 AGTTTATCATCTGCTTATAGTGAAGCTAGAATCATCCAGAATATAAATGCAACATTA      480
         ||||                                          ||||
Db  897 ATATTAACAACCACTAAAAAACTAAAAAAAAACATTATTAATAAATAAATCAACAAACT     838
```

Db   3461 TATTAAAAAATTTACTATAAATAAAAAGAAATATTTGCTCTAAAAAAAAATGTATAT 3520
          |||||
Qy   531 TAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAGC 590
          ||| ||||
Db   3521 TCATATGAGAATGGCAAAGAGAGAAAAAGTAACATAAAATTAGAAAAGAGCATTCAACA 3580
          ||| ||||
Qy   591 ACTAAAAACCACTTAGAACACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTA 650
          |||||
Db   3581 TTTTTTTATTTTGCTCAAGAAAAGAACACATATTGAAAATATTTAGTTCCCATTTTTT 3640
          ||| |||
Qy   651 TAATCAAATTCGCAATAATTAGTGGATCTATCAAGAAGCTAGTAGTTTAATAACTAA 710
          ||| ||||
Db   3641 TCAAAATAGAAAGATAATTATGGTAAACGATTTAATTAAACTAATCATCGAATAAAAAA 3700

Qy   711 AACACTAGATCCACTAAATGGGGGAACGCGTTTTTGATTCTAATGAGATTACTACAGTTAA 770
          |||||
Db   3701 TATAATATATAAAGCAAAACAGTGGAATTGTTAAAAATAAGGATAAGACATTTTTACATTT 3760
          |||||
Qy   771 TCGGAATATTAATAAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATT 830
          ||| ||||
Db   3761 AATCAAAAATAAAAGTAACAAAAATAACAATAACAAGAGAAGAACAAAAATAATTATAA 3820
          |||||
Qy   831 ATCTAAGTTTTATTTAAAAAAGTGATTCAAAATAATGAACAAAAGTTTGTAGGGACTTT 890
          ||||
Db   3821 CAATAATATATTAATAATAACAATAATAACAATAATTAATTAATTAATAAAAT 3880

Qy   891 TACAAAC 897
          ||||
Db   3881 TTAACAAC 3887

RESULT 11
US-09-790-988-1
; Sequence 1, Application US/0790988
; Patent No. 6632935
GENERAL INFORMATION:
 APPLICANT: SHIGENOBU, SHUJI
 APPLICANT: WATANABE, HIDEMI
 APPLICANT: HATTORI, MASAHIRA
 APPLICANT: SAKAKI, YOSHIYUKI
 TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
 FILE REFERENCE: 081356/0159
 CURRENT APPLICATION NUMBER: US/09/790,988
 CURRENT FILING DATE: 2001-02-23
 PRIOR APPLICATION NUMBER: JP2000-107160
 PRIOR FILING DATE: 2000-04-07
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
 LENGTH: 640681
 TYPE: DNA
 ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match   4.1%;   Score 54;   DB 4;   Length 640681;
Best Local Similarity   56.9%;   Pred. No. 0.03;
Matches   99;   Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy   16 TTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATA 75
          | |||
Db 604204 TTTATTTAAAATGCTGTTTATAAAAAAAAGAATAGCATTCTTTCAAATTTGTAAAAAATA 604263
          |||||
Qy   76 AGCCGTTAAATTAACTAACAAAAATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTA 135
          ||| |||
Db 604264 AGTTTTATAAATAATTTTAAAATTCATCTAATTTTTTTTAATTAAATTATATTCTAAA 604323
          ||| ||||
Qy   136 GTAATAAACGCTTATTTATTTTTATTTTTAGTCATCTTTTAAGATATAATAATATA 189
          |||||
Db 604324 ATAATAGAATAGACTTCTTTGCTATTTACAGGTAAATCACAAATTTTATTATA 604377

RESULT 12
US-10-204-708-82/c

──

Qy   481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
          ||| |||||
Db   837 TAAAAACAATAAATATAAAACCATCCATTATAAAACTATTACCAAAAAAAATTCCCGAT 778
Qy   541 AAAACGACTTTTGATAATGAACACCCAAATTAGTTGAAGCATACAAAACC 600
          ||| ||| |||
Db   777 CCAAATACCGAAAAAATTCTTAAACCTTAAATAAAAAAAATTCGAAACAAATCCATAAA 718
Qy   601 ACTTTAGAACAAC 613
          || ||||
Db   717 ATAAAATAAAAAAC 705

RESULT 10
US-09-417-485D-5
; Sequence 5, Application US/09417485D
; Patent No. 6541202
GENERAL INFORMATION:
 APPLICANT: Long, David M.
 APPLICANT: Metz, Anneke M.
 APPLICANT: Love, Ruschelle A.
 TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
 FILE REFERENCE: 47714-5009-US
 CURRENT APPLICATION NUMBER: US/09/417,485D
 CURRENT FILING DATE: 2002-06-14
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
 LENGTH: 10640
 TYPE: DNA
 ORGANISM: Plasmodium falciparum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (834)..(7385)
 OTHER INFORMATION: TERT gene
 FEATURE:
 NAME/KEY: unsure
 LOCATION: (1821)..(1837)
 OTHER INFORMATION: m at position 1821 = a or c; w at position 1837 =
 OTHER INFORMATION: a or t. Xaa (amino acid) at position 330 = Leu or
 OTHER INFORMATION: Ile; Xaa at position 335 = Asp or Gly.
US-09-417-485D-5

Query Match   4.2%;   Score 55;   DB 4;   Length 10640;
Best Local Similarity   42.7%;   Pred. No. 0.0078;
Matches   336;   Conservative 0; Mismatches 450; Indels 1; Gaps 1;

Qy   111 TTTTCTTATCAACCAAAATTCTCTAGTAATAAAACGCTTATTTATTTATTTTTAGTTCAT 170
          |||||
Db   3102 TTTTGTCTTTACCATCAAATTTTTAGGAATAAAAGTTAAAAAAAGAAATGAACCCAAA 3161
Qy   171 CTTTTAAGATATAAAATATGTTCAATGTGAAGAAAAAAGGTGAAAAAATAAAAACAAAT 230
          |||||
Db   3162 ATGGATTAATGTACAAAATATGTTCACAAAAAATGTCAATGTGAAGAAAAATAAAAACAAAT 3221
Qy   231 TATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTAT 290
          |||||
Db   3222 AAATATATATTTATTAAGAAAATGAAAAAAAGAGCACTAATAAATGTATTAATAATAAA 3281
Qy   291 TACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAGCGCGCGAATGGA 350
          |||||
Db   3282 TTTTCAAAAAATGTATCCCTAAAAAAAAAAAAATTTATATAACAATAACATCACACGTCAT 3341
Qy   351 GTTAACAGATCTAATCAATGCTAAAGCTAAAGCTTCACTACAAGACTATGCCAA 410
          |||||
Db   3342 AATATATATTTTATTAAAAGGATATGGAAAAAAATCAAAAACTAAACAATTTATTAT 3401
Qy   411 GATTGGAAGCTAGTTTATCATCTGCTTATGTGCTATATGTGAAGCTAAACAGTTAACAATAACCTTAA 470
          |||||
Db   3402 AAAAGTATAGATAATTTATACAAATTAAG-GAAATTAACAAAAAAGTGTTAGACCATA 3460
Qy   471 TGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAGCCATCAACCAAGC 530
          |||||

RESULT 13
US-09-601-198-39/c
; Sequence 39, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellson Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 1887
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-39

Query Match   4.1%;   Score 53.8; DB 4;   Length 5476;
Best Local Similarity 58.4%;   Pred. No. 0.012;
Matches 94;   Conservative 0;   Mismatches 67;   Indels 0;   Gaps 0;

Qy   52 TTCTAACAAATCCTAAATAATAGCCGTTAAATTAACTAAAAAATTAAAAAAATGGTT 111
          |||||||||||| ||| ||||||||||||
Db   5470 TTATATTTAAAAAAAAAAAAAAACTAAAAACATTACATAAAATATCAAAATATACGCTA 5411

Qy   112 TTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTTATTTTTAGTCATC 171
          ||||| |||||||||||||||||||||||||
Db   5410 ATCTTTTAACCCCAATAAACATATTTAAACATTATCCTATTTATTATTAATTTTCTTCATC 5351

Qy   172 TTTAAGAGATAATAAATATATCTTAATATTCTATGAATAAGAA 212
          |||||| |||||||||||||||
Db   5350 ATTTATTATATAATAATCAATTTTTAAATTCCTATAAATATAAAA 5310

Query Match   4.1%;   Score 53.6; DB 4;   Length 1887;
Best Local Similarity 43.4%;   Pred. No. 0.011;
Matches 349;   Conservative 0;   Mismatches 449;   Indels 6;   Gaps 2;

Qy   82 TAAATTAACTAAAAAATTAAAAAAATGTTTTCTTATCAACCAAAATTCTCTAGTAATA 141
          |||||| ||||
Db   1284 TTAGTTAATAATAACACAAAAGCTAATTTAAAAAACTATATGAAGATTTAAATTAATT 1225

Qy   142 AACGCTTATTTATTTTATTTTTAGTCATCTTTTAAGATACATATAATATATCTTAATATTCT 201
          ||| |||||||| |||||||||
Db   1224 TACACTAATTAAATAAATGATATTAATTAACAAGAACTAAAAAGCTTTACCTTTT 1165

Qy   202 ATGAATAGAAAAGAATCATCTTAAAGACTATTA----GTTTGTTAGGTACACACATCCT 256
          |||||||| |||||
Db   1164 ATTAAAAAATTAGATGATATTGTTAATGTTATTTACCAACAAAATTATAGT 1105

Qy   257 TTCTTAGCATTGGGATTTCTAGCTGTATGCTCATTACTAAAAAGATGCAAACCCAAATA 316
          |||||||||||||| |||||
Db   1104 TTTTTATGATTTGATCAATTAAAATTAACTAATTTAAAACAAAATCTAAACAAGAATATC 1045

Qy   317 ATGGCCAAACCCATTAGAAGCAGCGCCGAATGGAGTTAACAGATCTAATCAATGCTAAAG 376
          || |||
Db   1044 ACTAATATTGTTAATGTTAAAAAATCATTAGATATTGACAAATTAAATCTTAACTAAT 985

Qy   377 CGATGACATTAGCTTCACTACAGACTATGCCAAGATTGAA-GCTAGTTTATCATCTGCT 435
          ||||| ||||||||||||| |||
Db   984 AAATTATTAAAATGATTTACAAAATTATTTACAAGGTTTAAATACAATTGCATTAGATAGT 925

Qy   436 TATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACAGTTAGTTATTATTCAACAGATTACAAT 495
          ||||||||||||||| |||
Db   924 GAAAATAGTATTATAATTAATAAAGTTAGTATTATTAGTATGAATCAGCCATCAACCAAGCTAATACGGA 865

Qy   496 GCTAAAACTAATTTAGAATCAGCCTCAACCAAGCTAATATCCAAAAGACAAAAACAATAGTTCAAAA 555
          ||||| |||||||||||||
Db   864 CCTTACTATGTATTGAAAAACAACACAAATATCCAAAAACCACTTTAGAACACACGT 805

Qy   556 AATGAACACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAAACCCACTTTAGAACACACGT 615
          ||||||||||| |||
Db   804 GCTACATATAAATGGTTTAGAAAATTACAAATTGAAAATTACAATCAATTATTAGTTTATTT 745

Qy   616 GCTACTAACCTTGAAGGTTTGTCATCAACGCTTATATCAACTGCCAATAATTGCAATAATTTAGTG 675
          |||||| |||||||||||||
Db   744 AGCTTTAATACAAATAGCTTAAGCTAGTGTGCAAATAAATCTTAATAATGATTAACT 685

Qy   676 GATCTTACACAATAAAGCTAGTAGTTTAATAACTAGATCCACTAGTAGTGATCCACTAAAATGGGGA 735
          |||| |||||||||||||
Db   684 AAAAAATTTGAAGATGCTAATACTAAAATTAGTTCAATTGTTGATCAAATCAAACAAACA 625

Qy   736 ACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCCAGTTAATCGGAATATTAATAATACGTTATCA 795
          |||| |||||||||||||
Db   624 ATTCTATCAAACTTTAAAAAAACTTATTTATTTATTCAAAGATTTCATTAAACATCTAAAAGAA 565

Qy   796 ACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTCAAATAGTGTTTATTAAAAAAGTG 855
          |||||||| |||||||||
Db   564 TTTAATACTAAAGTAATAGTGAAAATGATCAAGCAAAGAAAACTAATAAAAACCCAAACA 505

Qy   856 ATTCAAAATAAATGAACAAAGTTTT 879
          |||||||| |||
Db   504 ATTCAAGATGTTTTAAATAGTTAT 481

RESULT 14
US-10-204-708-14/c
; Sequence 14, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708

; Sequence 82, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 82
; LENGTH: 5476
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-82

; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 14
; LENGTH: 6113
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-14

Query Match   4.1%;   Score 53.2;   DB 4;   Length 6113;
Best Local Similarity 46.3%;   Pred. No. 0.017;
Matches 253;   Conservative   0;   Mismatches 283;   Indels 10;   Gaps   2;

Qy   24 ATATCTTTGCTAAAAAACACAAAATCTTCAACAAAATCCTAAATAAATAAGCCGTTA 83
              | ||||| | |||| |||| ||||||| | | || | |||
Db   4095 AAATTATACATTTTAAACATAAAACAATTTTTTTATCAATTATACTTCAATAAAACTCTC 4036

Qy   84 AATTAACTAAAAAATTAAAAAAATGGTTTTTCCTTATCAACCAAAATTCTCTAGTAATAAA 143
              || ||| || | || ||| | | || | ||| ||||||| ||
Db   4035 AAAATACAATTACAATTACAAATTATAAAAAAATATACAAATTAATAAAAATAAAAATAAA 3976

Qy   144 CGCTTATTTATTTTATTTTAGTCATCTTTAAGATATAAATATATCTTAATATTCTAT 203
              || |||| | ||| | ||| ||| ||| | || | |||| |||
Db   3975 ATATTCATATTATTATATAAAATAAACATAACTATTAAATTAAACTTAATATATATTATAT 3916

Qy   204 GAATAAGAAAGAATCATCTTAAAGACTTAGTTGTTGGTTACAACATCCTTTCTTAG 263
              |||||| || || ||| | ||| || || | ||| || |||
Db   3915 CCCTATATAAAATAATACAAAACTAACTATTAATAATAATATTA-------TTTATTAT 3864

Qy   264 CATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAAATGGCCA 323
              ||| |||||| ||| |||| |||| | | |||| | | ||| ||| |
Db   3863 TTTTTAAACTTATTTTTTTTTTTTTTTTTAAAATAAAATCTCACTCATCACCCAAACTA 3804

Qy   324 AACCCAATTAGAGCAGCGCGAATGAGAGTTAACAGATCTAATCAATGCTAAAGCGATGAC 383
              |||||| |||| |||| | |||| ||| || || | ||||||| | |||
Db   3803 AAATACAATAAATAATAATCTCGACTCACTACAACATCCGCCTCCCAAATTCAAATAATTTT 3744

Qy   384 ATTAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGA 443
              ||| | ||| | | || || | | | | || || || || |
Db   3743 CCTACCTTAACCTCCCGACTAACTAACTAAAATTAACACAAAC--GCGTACCACCATACCCACTAA 3686

Qy   444 AGCTGAAACAGTTAACACATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAC 503
              ||| |||| | |||| | | || | || ||||| | |||
Db   3685 TTTTTATATTTTTAATAAAATAAAATTTCACCATATTAACCAAACTAATCTAAAACTCC 3626

Qy   504 TAATTTAGAATCAGCCATCAACAAGCTAATACGGAATAAACGACTTTTGATAAATGAACA 563
              |||||| | | || ||| | || || || || || | || ||| ||
Db   3625 TAATTCAAAATAATCCACCCCACCTCGACCTCCCAAAATAATAAAATTACAAACATAAACC 3566

Qy   564 CCCAAA 569
              ||| |
Db   3565 ACCGAA 3560

RESULT 15
US-10-204-708-12/c
; Sequence 12, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 12
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-12

Query Match   4.1%;   Score 53.2;   DB 4;   Length 6317;
Best Local Similarity 49.3%;   Pred. No. 0.017;
Matches 139;   Conservative   0;   Mismatches 143;   Indels   0;   Gaps   0;

Qy   34 TTAAAAAACACAAAATCTCTAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAA 93
              ||| ||| ||| ||| | ||| | ||| | ||| | || || ||
Db   5042 TTTATCAAAAAAAAAATAATCTCCTTAACAAAAAAAACCAACATTAACAATA 4983

Qy   94 AAAATTAAAAAAATGGTTTTTCTTATCAACCAAAATCCTAGTAATAAACGCTTATTTA 153
              ||| |||||| | || || | | | ||| || ||| |||| | ||
Db   4982 AAAAATAAAAATAAAAAAAAATTACATTATCAAAATCATAAAACTAAAATCTA 4923

Qy   154 TTTTTATTTTTAGTCATCTTTAAGATATAAATATATCTTCTATGAATAAGAAA 213
              ||||| |||| | |||| || ||| |||| ||| || |||| ||||||
Db   4922 TTTATCTTAACAAATATTATTATTATAATTACTACTTTTTAAAAAAATTCA 4863

Qy   214 AGAATCATCTTAAAGACTATTAGTTGTAGGTACAACATCCTTTCTTAGCATTGGGATT 273
              |||| ||| | |||| ||| || || |||| | || ||| || || || ||||
Db   4862 TTTTTCATATAATTTTCATAACAAAATATTATATAATTATTACTAACATTTTTCACT 4803

Qy   274 TCTAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAAT 315
              || | | ||| | |||| |||| | || | |||| |||
Db   4802 TTTAACTTTTACAAAATAAATAAATAATATTTTTTACAAAAAT 4761

Search completed: May 5, 2004, 20:04:08
Job time : 106.495 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    May 5, 2004, 17:07:04 ; Search time 514.265 Seconds
                                    (without alignments)
                            11489.234 Million cell updates/sec

Title:              US-09-901-572A-1
Perfect score:      1306
Sequence:           1 aaaaacatcagattgttaat........taaaatcgttttatcaggtt 1306

Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0
Searched:           2937390 seqs, 226062796 residues

Total number of hits satisfying chosen parameters:        5874780

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications NA:*
           1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
           2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
           3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
           4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
           5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
           6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
           7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
           8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
           9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
          10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
          11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
          12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
          13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
          14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
          15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
          16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
          17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
          18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
          19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	*Query Match	Length	DB	ID	Description
1	1306	100.0	1306	10	US-09-901-572A-1	Sequence 1, Appli
2	1028	78.7	1152	15	US-10-131-591A-1	Sequence 1, Appli
3	1024.2	78.4	1371	9	US-09-147-052-1	Sequence 1, Appli
4	1024.2	78.4	3261	9	US-09-147-052-3	Sequence 3, Appli
5	921.6	70.6	1082	15	US-10-131-591A-24	Sequence 24, Appl
c 6	73.4	5.6	4985	15	US-10-094-240-10	Sequence 10, Appl
c 7	73.4	5.6	4985	15	US-10-056-405-10	Sequence 10, Appl
c 8	73	5.6	6292	13	US-10-221-714A-461	Sequence 461, App
c 9	69.6	5.3	7892	17	US-10-257-166-138	Sequence 138, App
c 10	68.6	5.3	8136	15	US-10-311-455-528	Sequence 528, App
c 11	68.6	5.3	8136	17	US-10-257-166-38	Sequence 38, Appl
c 12	68.4	5.2	11836	15	US-10-239-676-102	Sequence 102, App
c 13	68.4	5.2	11836	15	US-10-240-453-114	Sequence 114, App
c 14	68.2	5.2	8392	15	US-10-311-455-1463	Sequence 1463, Ap

c 15	68	5.2	12405	13	US-10-221-613-101	Sequence 101, App
c 16	68	5.2	12405	15	US-10-239-676-35	Sequence 35, Appl
c 17	68	5.2	12405	15	US-10-240-453-43	Sequence 43, Appl
c 18	68	5.2	3673778	15	US-10-312-841-1	Sequence 1, Appli
c 19	67	5.1	8076	17	US-10-257-166-36	Sequence 36, Appl
c 20	66.4	5.1	6669	15	US-10-204-708-6	Sequence 6, Appli
c 21	66.4	5.1	6669	17	US-10-311-455-166	Sequence 166, App
c 22	66.4	5.1	6669	15	US-10-240-589C-6	Sequence 6, Appli
c 23	65.8	5.0	6264	13	US-10-221-714A-333	Sequence 333, App
c 24	65.8	5.0	6264	15	US-10-240-452-67	Sequence 67, Appl
c 25	65.6	5.0	3057	15	US-10-349-680-148	Sequence 148, App
c 26	65.6	5.0	5163	15	US-10-311-455-1221	Sequence 1221, Ap
c 27	65.4	5.0	8771	15	US-10-311-455-1797	Sequence 1797, Ap
c 28	64.8	5.0	7351	15	US-10-311-455-1	Sequence 1, Appli
c 29	64.8	5.0	18512	15	US-10-311-455-950	Sequence 950, App
c 30	64.8	5.0	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 31	64.2	4.9	17594	15	US-10-311-455-1999	Sequence 1999, Ap
c 32	63.2	4.8	16766	15	US-10-311-455-2130	Sequence 2130, Ap
c 33	63	4.8	9254	13	US-10-221-714A-114	Sequence 114, App
c 34	63	4.8	9254	15	US-10-239-676-86	Sequence 86, Appl
c 35	63	4.8	9254	15	US-10-240-453-92	Sequence 92, Appl
c 36	62.6	4.8	5883	15	US-10-311-455-326	Sequence 326, App
c 37	62.6	4.8	5883	15	US-10-240-485-26	Sequence 26, Appl
c 38	62.6	4.8	8591	13	US-10-221-714A-291	Sequence 291, App
c 39	62.2	4.8	5678	15	US-10-311-455-1111	Sequence 1111, Ap
c 40	62.2	4.8	6061	15	US-10-311-455-114	Sequence 114, App
c 41	62.2	4.8	9728	15	US-10-311-455-1876	Sequence 1876, Ap
c 42	62.2	4.8	9728	17	US-10-257-166-174	Sequence 174, App
c 43	62	4.7	5520	15	US-10-311-455-1492	Sequence 1492, Ap
c 44	61.4	4.7	6012	13	US-10-221-613-235	Sequence 235, App
c 45	61.4	4.7	7001	15	US-10-172-086-60	Sequence 60, Appl

ALIGNMENTS

```
RESULT 1
US-09-901-572A-1
; Sequence 1, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
;  APPLICANT: Nippon Zeon Co., Ltd.,
;  TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
;  FILE REFERENCE: J209
;  CURRENT APPLICATION NUMBER: US/09/901,572A
;  CURRENT FILING DATE:  2003-03-11
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
;  LENGTH: 1306
;  TYPE: DNA
;  ORGANISM: Mycoplasma gallisepticum
;  FEATURE:
;  OTHER INFORMATION: TTM-1 gene
US-09-901-572A-1

Query Match             100.0%;  Score 1306;  DB 10;  Length 1306;
Best Local Similarity   100.0%;  Pred. No. 7.1e-229;
Matches 1306;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAAACATCAGATTGTTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAA 60

Qy     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAAAAAATGGTTTTCTTATC 120

Qy    121 AACCAAAATTCCTAGTAATAAACGCTTATTTTAATTTTATTTTTAGTCATCTTTTAAGAT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AACCAAAATTCCTAGTAATAAACGCTTATTTTAATTTTTATTTTTAGTCATCTTTTAAGAT 180

Qy    181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
```

Db   1261 CCAGCAGTTGATGAGATTAAGAGTTGCTAAAATCGTTTATCAGGTT 1306

RESULT 2

US-10-131-591A-1

; Sequence 1, Application US/10131591A
; Publication No. US2003059799A1
; GENERAL INFORMATION:
;   APPLICANT: Nippon Zeon Co., Ltd.
;   TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
;   FILE REFERENCE: J209
;   CURRENT APPLICATION NUMBER: US/10/131,591A
;   CURRENT FILING DATE: 2002-08-15
;   NUMBER OF SEQ ID NOS: 79
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Mycoplasma gallisepticum
;   FEATURE:
;   OTHER INFORMATION: TTM-1 gene (after EcoRI)

US-10-131-591A-1

Query Match   78.7%;   Score 1028;   DB 15;   Length 1152;
Best Local Similarity   100.0%;   Pred. No. 4.1e-178;
Matches 1028;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy  279 CTGTATGTCTATTACTAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGC 338
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6 CTGTATGTCTATTACTAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAGC 65

Qy  339 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCTAGCTTCACTCACCA 398
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCTAGCTTCACTCACA 125

Qy  399 AGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAAACAGTTAA 458
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126 AGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAAACAGTTAA 185

Qy  459 CAATAACCTTAATGCAACATTAGCAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGC 518
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  186 CAATAACCTTAATGCAACATTAGCAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGC 245

Qy  519 CATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGA 578
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  246 CATCAACCAAGCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGA 305

Qy  579 AGCATACAAAGACTAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 638
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  306 AGCATACAAAGACTAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTC 365

Qy  639 ATCAACTGCTATAATAACACTAGATCCACTAAATGGGGAACCGCTTTTAGATTCTAATGAGAT 698
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  366 ATCAACTGCTATAATAACACTAGATCCACTAAATGGGGAACCGCTTTTAGATTCTAATGAGAT 425

Qy  699 TTTAATAACTAAAAACACTAGATCCACTAAATGGGGAACCGGAACAAACACGGAAG 758
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  426 TTTAATAACTAAAAACACTAGATCCACTAAATGGGGAACCGGAACAAACACGGAAG 485

Qy  759 TACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATTAATGAACAAAAGAACTAA 818
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  486 TACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATTAATGAACAAAAGAACTAA 545

Qy  819 TGCTGATGCATTATCATAGTTTATTAAAAAGTGATTCAAATAATGACAAAGTTT 878
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546 TGCTGATGCATTATCATAGTTTATTAAAAAGTGATTCAAATAATGACAAAGTTT 605

Qy  879 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAG 938
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  606 TGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAG 665

Qy  939 TGCTGATGTAACCACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGA 998
```

Db   181 ATAAATATATCTTAATATTCTATGAATAGAAAGAATCATCTTAAAGACTATTAGTTTG 240

Qy   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAA 300
Db   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAA 300

Qy   301 GATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGAGTTAACAGAT 360
Db   301 GATGCAAACCCAATAATGGCCAAACCCAATTAGAAGCAGCGCGGAATGGAGTTAACAGAT 360

Qy   361 CTAATCAATGCTAAAGCGATGACTACTAGCTTCACTACAGACATGCCAAGATTGAAGCT 420
Db   361 CTAATCAATGCTAAAGCGATGACTACTAGCTTCACTACAGACATGCCAAGATTGAAGCT 420

Qy   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480
Db   421 AGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTA 480

Qy   481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540
Db   481 GAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGAT 540

Qy   541 AAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGACTAAAAACC 600
Db   541 AAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGACTAAAAACC 600

Qy   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTATAATAATCAAATT 660
Db   601 ACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTATAATAATCAAATT 660

Qy   661 CGCAATAATTTAGTGGATCTATACAGCTAGTAGTTAATAAAGCTAGAGATCAAAACACTAGAT 720
Db   661 CGCAATAATTTAGTGGATCTATACAGCTAGTAGTTAATAAAGCTAGAGATCAAAACACTAGAT 720

Qy   721 CCACTAAATGGGGAACCGCTTTTAGATTCTAATGAGAATTACTACAGTTAATCGGAATATT 780
Db   721 CCACTAAATGGGGAACCGCTTTTAGATTCTAATGAGAATTACTACAGTTAATCGGAATATT 780

Qy   781 AATAATACGTTATCAACTATTATTAATGAACAAAAGAACTAATGCTGATGCATTATCTAATAGT 840
Db   781 AATAATACGTTATCAACTATTATTAATGAACAAAAGAACTAATGCTGATGCATTATCTAATAGT 840

Qy   841 TTTATTAAAAAGTGATTCAAAATAATGACAAAGTTTTGTAGGGACTTTTACAAACGCT 900
Db   841 TTTATTAAAAAGTGATTCAAAATAATGACAAAGTTTTGTAGGGACTTTTACAAACGCT 900

Qy   901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACCACCCGTCAAT 960
Db   901 AATGTTCAACCTTCAAACTACAGTTTGTTGCTTTTAGTGCTGATGTAACCACCCGTCAAT 960

Qy   961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020
Db   961 TATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGAACCTTCAAGTAGAATTCTTGCA 1020

Qy   1021 AACACGAATAGTATCACAGATGTTCTTGGATTTAGCTGGAACAAACACGAAG 1080
Db   1021 AACACGAATAGTATCACAGATGTTCTTGGATTTAGCTGGAACAAACACGAAG 1080

Qy   1081 TACCAATTTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG 1140
Db   1081 TACCAATTTAGTTTAGCAACTATGGTCCATCAACTGGTTATTTATATTCCCTTATAAG 1140

Qy   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGGAAAT 1200
Db   1141 TTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTACAATACAAATAAATAATGGAAAT 1200

Qy   1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATAACTACAGCTAATCCAACT 1260
Db   1201 GTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAAATAATAACTACAGCTAATCCAACT 1260

Qy   1261 CCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTTATCAGGTT 1306

Db   666 TGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATGA 725

Qy   999 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAG 1058
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 ACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTTATAG 785

Qy   1059 TTTAGTCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATCGTCCATCAACTGG 1118
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 TTTAGTCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATCGTCCATCAACTGG 845

Qy   1119 TTATTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCGTAATAACGTTGGATTACA 1178
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 TTATTTTATATTTCCCTTATAAGTTGGTTAAAGCAGCTGATGCGTAATAACGTTGGATTACA 905

Qy   1179 ATACAAAATTAAATAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 1238
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 ATACAAATTAAATAAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTGCAAA 965

Qy   1239 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAAATCGTTTT 1298
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 TAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAGTTGCTAAAATCGTTTT 1025

Qy   1299 ATCAGGTT 1306
          ||||||||
Db   1026 ATCAGGTT 1033

RESULT 3
US-09-147-052-1
; Sequence 1, Application US/09147052
; Patent No. US2001001435A1
; GENERAL INFORMATION:
;   APPLICANT: SAITOH, Shuji
;   APPLICANT: TSUZAKI, Yoshinari
;   APPLICANT: YANAGIDA, No. US2001001435A1oru
;   TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
;   TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
;   FILE REFERENCE: 981167
;   CURRENT APPLICATION NUMBER: US/09/147,052
;   CURRENT FILING DATE: 1999-04-05
;   PRIOR APPLICATION NUMBER: JP 08-103548
;   PRIOR FILING DATE: 1996-03-29
;   PRIOR APPLICATION NUMBER: PCT/JP97/01084
;   PRIOR FILING DATE: 1997-03-28
;   NUMBER OF SEQ ID NOS: 9
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1371
;   TYPE: DNA
;   ORGANISM: hybrid
US-09-147-052-1

Query Match   78.4%;   Score 1024.2;   DB 9;   Length 1371;
Best Local Similarity 99.7%;   Pred. No. 2.2e-177;
Matches 1026;   Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

Qy   278 GCTGTATGTCTATTACTAAAAAAGATGCAAACCCAATPAATGGCCAAACCCAATTAGAAG 337
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 GCTGTATGTCTATTACTAAAAAAGATGCAAACCCAATAATGGCCAAACCCAATTAGAAG 247

Qy   338 CAGCGCGAATGGAGTGAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 397
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   248 CAGCGCGAATGGAGTGAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 307

Qy   398 AAGACTATGCCAAGATTGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTGAAACAGTA 457
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   308 AAGACTATGCCAAGATTGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA 367

Qy   458 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAG 517
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   368 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCAG 427

Qy   518 CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTAGTTG 577

Db   428 CCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTAGTTG 487

Qy   578 AAGCATACAAAGCACTAAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT 637
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   488 AAGCATACAAAGCACTAAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT 547

Qy   638 CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTA 697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   548 CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTA 607

Qy   698 GTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 757
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   608 GTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGA 667

Qy   758 TTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 817
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   668 TTACTACAGCTAATAAGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 727

Qy   818 ATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTT 877
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   728 ATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTT 787

Qy   878 TTGTAGGGACTTTACAAACGTAATGTCAACCTTCAAACTACACAGTTTTTGTTGCTTTTA 937
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   788 TTGTAGGGACTTTACAAACCTAATGTCAACCTTCAAACTACACAGTTTTTGTTGCTTTTA 847

Qy   938 GTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATG 997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   848 GTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAATGGTGATG 907

Qy   998 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTGGATTTATA 1057
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   908 AACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTGGATTTATA 967

Qy   1058 GTTTAGTCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATCGTCCATCAACTG 1117
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   968 GTTTAGTCTGGAACAAACACGAGTACCAATTTAGTTTTAGCAACTATCGTCCATCAACTG 1027

Qy   1118 GTTATTTATATTTCCCTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATTAACGTTGGATTAC 1177
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1028 GTTATTTATATTTCCCTTATATTTCCCTTATAGTTGGTTAAAGCAGCTGATGATGTAATAACGTTGGATTAC 1087

Qy   1178 AATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGGAGATGGTGCTAAAATCGTTT 1237
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1088 AATACAAATTAAATAATGGAAATGTTCAACAAGTTGAGTTGGAGATGGTGCTAAAATCGTT 1147

Qy   1238 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1148 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAATCGTTT 1207

Qy   1298 TATCAGGTT 1306
          |||||||||
Db   1208 TATCAGGTT 1216

RESULT 4
US-09-147-052-3
; Sequence 3, Application US/09147052
; Patent No. US2001001435A1
; GENERAL INFORMATION:
;   APPLICANT: SAITOH, Shuji
;   APPLICANT: TSUZAKI, Yoshinari
;   APPLICANT: YANAGIDA, No. US2001001435A1oru
;   TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
;   TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE
;   FILE REFERENCE: 981167
;   CURRENT APPLICATION NUMBER: US/09/147,052
;   CURRENT FILING DATE: 1999-04-05
;   PRIOR APPLICATION NUMBER: JP 08-103548
;   PRIOR FILING DATE: 1996-03-29
;   PRIOR APPLICATION NUMBER: PCT/JP97/01084
;   PRIOR FILING DATE: 1997-03-28

Query Match   78.4%;   Score 1024.2; DB 9;   Length 3261;
Best Local Similarity   99.7%; Pred. No. 3.1e-177;
Matches 1026; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

Qy   278 GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 337
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2078 GCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAG 2137

Qy   338 CAGCGCGAATGGAGTTAACAGATCTAATCAATCGATGACATTAGCTTCACTAC 397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2138 CAGCGCGAATGGAGTTAACAGATCTAATCAATCGATGACATTAGCTTCACTAC 2197

Qy   398 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA 457
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2198 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTA 2257

Qy   458 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAAATCAG 517
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2258 ACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAAATCAG 2317

Qy   518 CCATCAACCAAGCTAATACCGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG 577
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2318 CCATCAACCAAGCTAATACCGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGTTG 2377

Qy   578 AAGCATACAAAGCACTAAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGT 637
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2378 AAGCATACAAAGCACTAAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGT 2437

Qy   638 CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGTAGTA 697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2438 CATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCTATACAATAAAGTAGTA 2497

Qy   698 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGGAACGCCTTTTAGATTCTAATGAGA 757
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2498 GTTTAATAACTAAAACACTAGATCCACTAAATGGGGGAACGCCTTTTAGATTCTAATGAGA 2557

Qy   758 TTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 817
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2558 TTACTACAGCTAATAAGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGACTA 2617

Qy   818 ATGCTGATGCATTATCTAATAGTTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTT 877
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2618 ATGCTGATGCATTATCTAATAGTTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTT 2677

Qy   878 TTGTAGGGACTTTACACCCGTCAATTATAAATAATGCAAGAAGGACCGTTTGGAAATGGTGATG 937
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2678 TTGTAGGGACTTTACAACCCGTCAATTATAAATAATGCAAGAAGGACCGTTTGGAAATGGTGATG 2737

Qy   938 GTGCTGATGTAACACACCCGTCAATTATAAATAAATGCGAATAGTATCCACAGAATGTTCTTGGATTTATA 997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2738 GTGCTGATGTAACACCCGTCAATTCAATTATAAATAATGCGAATAGTATCCACAGAATGTTCTTGGATTTATA 2797

Qy   998 AACCTTCAAGTAGAATTCTTGCAAACACGAATGAATATCCACAGAATGTTTCTTGGATTTATA 1057
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2798 AACCTTCAAGTAGAATTCTTGCAAACACGAATGAATATCCACAGAATGTTTCTTGGATTTATA 2857

Qy 1058 GTTTAGCTGGAACAAAACGAAGTACCAATTTAGTTTTAGCAACTATGGTCCATCAACTG 1117
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2858 GTTTAGCTGGAACAAAACGAAGTACCAATTTTAGTTTTAGCAACTATGGTCCATCAACTG 2917

Qy 1118 GTTATTTATATTCCCTTATAGTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC 1177
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2918 GTTATTTATATTTCCCTTATAGTGGTTAAAGCAGCTGATGCTAATAACGTTGGATTAC 2977

Qy 1178 AATACAAATTAAATAATGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 1237

Db 2978 AATACAAATTAAATAATGAAATGTTCAACAAGTTGAGTTTGCCACTTCAACTAGTGCAA 3037

Qy 1238 ATATATTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTT 1297
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3038 ATAATACTACAGCTAATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAAAATCGTTT 3097

Qy 1298 TATCAGGTT 1306
          |||||||||
Db 3098 TATCAGGTT 3106

Query Match   70.6%;   Score 921.6; DB 15;   Length 1082;
Best Local Similarity   98.0%; Pred. No. 1.1e-158;
Matches 933; Conservative   0; Mismatches 19; Indels   0; Gaps   0;

Qy   355 ACAGATCTAATCAATGCTAAAGCTAGTCTCACTACAAGACTATGCCAAGATT 414
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACAGATCTAATCAATGCTAAAGCTAGTCTCACTACAAGACTATGCCAAGATT 60

Qy   415 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCA 474
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAACCTGCAGCA 120

Qy   475 ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCGCATCAACCAAGCTAAT 534
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCGCATCAACCAAGCTAAT 180

Qy   535 ACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 594
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 ACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTA 240

Qy   595 AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATATAAT 654
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATATAAT 300

Qy   655 CAAATTCGCAATAATTAGTGGGGACAACGCGTTTTGGATCTATACAAGTCAGTTAATCGG 714
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CAAATTCGCAATAATTAGTGGGGACAACGCGTTTTGGATCTATACAAGTCAGTTAATCGG 360

Qy   715 CTAGATCCACTAAATGGGGGAACCGTTTAGATTCTCAATGAAGCTAGTTTAACAATACTACAGCTAATAAG 774
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTAGATCCACTAAATGGGGGAACCGTTTAGATTCTCAATGAAGCTAGTTTAACAATACTACAGCTAATAAG 420

Qy   775 AATAATAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCT 834
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AATAATCCAGAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCT 480

Qy   835 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA 894
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAGTTTTGTAGGGACTTTTACA 540

RESULT 6
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US2003082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT APPLICATION NUMBER: US/10/094,240
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match   5.6%;   Score 73.4;   DB 15;   Length 4985;
Best Local Similarity 43.3%;   Pred. No. 0.0018;
Matches 392;   Conservative   0;   Mismatches 511;   Indels 2;   Gaps 1;

RESULT 7
US-10-056-405-10/c
; Sequence 10, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-056-405-10

Query Match   5.6%;   Score 73.4;   DB 15;   Length 4985;
Best Local Similarity 43.3%;   Pred. No. 0.0018;
Matches 392;   Conservative   0;   Mismatches 511;   Indels 2;   Gaps 1;

us-09-901-5/2a-1.may2004.rapb

; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
SEQ ID NO 461
; LENGTH: 6292
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-461

Query Match   5.6%;   Score 73; DB 13; Length 6292;
Best Local Similarity 43.9%;   Pred. No. 0.0023;
Matches 368; Conservative 0; Mismatches 465; Indels 6; Gaps 1;

Qy   36 AAAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAA 95
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   4184 AAAAAAAACAAAAAAAAAAAAAAACGAACAAAAAAAAAAAAAATTTTATAAAAAAACAAT 4125

Qy   96 AATTAAAAAAAATGGTTTTCTTATCAACCAAAATTCCTAGTAATAAACCGCTTATTTATT 155
          ||||||| ||||||||||
Db   4124 AAATTAATTAAATAATAATTAAATATAATAATTTAAATTTAAATAATAACCAATAAAAAAT 4065

Qy   156 TTTATTTTTAGTCATCTTTAAGATATAAATATTCTTAATATTCTATGATAAGAAGAAAAG 215
          ||||||| |||||||
Db   4064 AAAAAAAATAACGAACCGAAAAAAAAAAAAATTTTACTAAAAAACGATAAAAACAAAAAA 4005

Qy   216 AATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTC 275
          ||||| |||||||
Db   4004 AAATATATAAAAACAATAACTAAAAAATATTTAAATAAACGAAAAAAAAACCCAATTAGA 3945

Qy   276 TAGTCGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAATAAAAACTAAA 335
          |||| |||||||
Db   3944 AATAACAAATAAAAAATAAAATATACCAAAAAAAATAAAAAAAAAAAAAACTAAA 3885

Qy   336 AGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACT 395
          ||| |||||||
Db   3884 AAAAAAACGATAAAAAAATATTTAAAATAATAAATAAAACAAAAAAAAAATAACAAACA 3825

Qy   396 ACAAGACTATGCCAAGATTGAAGCTAGTCAGTTTATCATCTCGCTTATAGTGTGAAGCAGT 455
          ||| |||||||
Db   3824 AAATAAAAAAAAAAAAAGCAACGAAAAAAAAACTTTATAAAAAATAACAAA 3765

Qy   456 TAACAATAACCTTAATGCAACATTAGAACACTAAAAATGGCTAAAACTAATTAGAATC 515
          | |||||||
Db   3764 AAAAAATTAAAAAAATAAATAAATTAA------AAAAACGTAATAATAACAAAAATAAAAAA 3711

Qy   516 AGCCATCAACCAAGCTAATACGCGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGT 575
          ||| |||||||
Db   3710 AAAACTTTACAAAAAAACGACAAAAAAATATTTTTAAATAATAAATAAAAAAAAAAAT 3651

Qy   576 TGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTT 635
          ||| |||||||
Db   3650 AACAAAAAAAAAAAAAAAAACGATAAAAAAAACAATTTTTAAATAAACGAAAAAAAAA 3591

Qy   636 GTCATCAACTGCTTATAATCAAATTGCAATAATTTAGTGGATCTATACAATAAAGCTAG 695
          ||| |||||||
Db   3590 AAAAAAAATAACGAACACACAAAAATAAAAAAAAAAAAAACTTTAAAAAAAAAAATAAAATAAATAT 3531

Db   3163 TATTGTGTATTACTACAGTTACTCGGAATTAATTATTTACTATTATTTACTTTATTA 3104
          ||| ||| ||||||
Qy   170 TCTTTTAAGATATAAATATATCTTAATATTCCATGATAAGAAGAAAGAATCATCTTAAAGA 229
          ||| |||
Db   3103 TTTTTTATTTATGGATTACTGGAATAATTTAGAGAATAAAAAAAACAATAATAATACACA 3044

Qy   230 CTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTTCAGTCGTGATGTCTA 289
          |||||||
Db   3043 TAATAAGAATAACAACAACAACAATAAGAACAACAACACAAAACAATAATAATAAGAATAA 2984

Qy   290 TTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGCGAATGG 349
          ||| |||
Db   2983 TAATAATACAATAGCAATAATATAATACAATAATAGCAATAATAATAATAATAATAATA 2924

Qy   350 AGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGACTATGCCA 409
          ||||||
Db   2923 TAATAATAATAATAATAATAATAATATTAAAAAAAAAAATAATTAATATAATGAATAGTA 2864

Qy   410 AGATTGAAGCTAGTTTATCATCTCGTTAGTGAAGCTGAAACAGTTAACAATAACCTTA 469
          ||| |||||
Db   2863 ATAATAATAATAATAATATTAAAAAAATAATTAATAATAATAATAATAATAATAATAATA 2804

Qy   470 ATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAG 529
          ||| |||
Db   2803 AGAATAATAATAATATAATAATAATAATAATAATAATAATCAATAATAATCAATAATAAAA 2744

Qy   530 CTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAG 589
          ||| |||||
Db   2743 TTATTATTATGATAATAATGTATAATAATGAATAACATAATAATAA--TAATTAATAATAAT 2686

Qy   590 CACTAAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTT 649
          ||| |||
Db   2685 AAATTAATTAATTAATAATAATAATAATAATTATTATTGTTATAATAATAATAATAATA 2626

Qy   650 ATAATCAAATTCGCAATAATTTAGTGGATCTATCACAATAAAGCTAGTTTAATAACTA 709
          ||| |||
Db   2625 ATAATAATAATAATAATCATAATTATTAATAATAATAATAATAATAATAATAATAATA 2566

Qy   710 AAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGAGATTACTACAGTTA 769
          ||| |||||
Db   2565 ATAATAATAATAATAATAATAATAATAATAAGTAATAATATAATAATAATAAATGTA 2506

Qy   770 ATCGGAATTAATAATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATGCAT 829
          ||| |||||
Db   2505 ATAATAATAATAATAATAATAATAATAATTAATAATAATAATATAATATAATAATA 2446

Qy   830 TATCTAATAGTTTTATTTAAAAGTCATTCAAAATAATGAACAACAAGTTTTGTAGGGACTTT 889
          ||| |||||
Db   2445 ATAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 2386

Qy   890 TTACAAACGCTAATGTTCAACCTTCAACTACAGTTTTGTGCTTTTAGTGCTGATGTAA 949
          ||| |||||
Db   2385 ATAATAATAATAATAATAATAATAATAATAATTCTGTATTTGTATTTTGGTTCCTGTAA 2326

Qy   950 CACCCGTCAATTATAAAATATGCAAGAAGGACCGTTGAAATGGTGATGAACCTTCAGTA 1009
          ||| |||
Db   2325 CTAAATTGGTTACTAAAAATGGTTTATCTTTCTCTTAAAATAAAAATCAATCCCTAAAGGA 2266

Qy   1010 GAATT 1014
          |||||
Db   2265 AGCTT 2261

RESULT 8
US-10-221-714A-461/c
; Sequence 461, Application US/10221714A
; Publication No US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with

```
Qy  696  TAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGGCTTTTAGATTCTAATGA  755
              ||||| ||||   |||| |    |||    || |||      |||||    ||||||
Db 3530  TTTTAAAAATAAAAAACAAATAAAAAAAAAAAAATAATAAAAAAAAAAAATA          3471

Qy  756  GATTACTACAGTTAATCGGAATATTAATATATACGTTATCAACATTATTAATGAACAAAGAC  815
              ||| |  ||| ||    | ||||  |||| |   || ||  |||| || |||||||||
Db 3470  CCAAAAAAAAAAAAAAACACGATAAAAACAATTTAAAATAATAATAAATAAAAAAAACG     3411

Qy  816  TAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAA     874
           |   || |||     ||||    |||||| || |||||        || |||||||||
Db 3410  AAAATAAACGTATATATATATATAAAAAACTTTACAAAAAAAAAAAAAAAAAAAATAA     3352
```

RESULT 9
US-10-257-166-138/c
; Sequence 138, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
;   DE 10032529.7
;   DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
;   2000-06-30
;   2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 138
; LENGTH: 7892
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-138

```
Query Match            5.3%;  Score 69.6;  DB 17;  Length 7892;
Best Local Similarity 46.0%;  Pred. No. 0.011;
Matches  309;  Conservative   0;  Mismatches  359;  Indels   4;  Gaps   2;
```

```
Qy   36  AAAAAAACACAAAATCTTCTACAACAAAAATCCTAAATAAATAAGAGCCGTAAAATAAATTAACTAAAAA  95
              |||  |||   ||||| |||  |   ||| |||| |       |   ||||    ||||  |||||
Db  883  ACAATAAATAAAAAAATACTTCATTAATAAAAAAAAAAATAATCCAAAAAATAAACTTTAA        824

Qy   96  AATTAAAAAAAATGGTTTTTCTTATCAACCAAAATCTCTAGTAAAAACCGCTTATTTATT         155
              ||||||    ||     |  ||||   |||||     ||||        |||||  |
Db  823  AATAAAAAAAACAAAAAACAAAAAATAATAAATATTTTTAATTAAAACTAAATAAAATA         764

Qy  156  TTTATTTTTAGTGTCATCTTTTAGATATAAAGATGATAATATCTTAATATTCTATGAATAAGAGAAAG  215
           || |  ||| |       ||||||||      | ||||||| | |||   ||    |||| |||||
Db  763  TCAACTCATAAAAAAATAATATTCCCAAAAAAAAAATCATAAAATTACAAAAATATAAAATTTAAAAT  704

Qy  216  AATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTC         275
           | || |  |||   ||      ||      ||  | |   |     ||    |||||||
Db  703  ACATAACAATAACAATAAAAAAAATCAAAATAAAAATAAAAAAAAA--AACTTAAAAATTCT       647

Qy  276  TAGCTGTATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCCAATTAGA        335
              || || |   ||||       ||  | ||||||   ||||| |    |   |||||||
Db  646  AAAATCATTACATTATCCTAAAAAAAAAAAAATACAATAATTAATCAAAAATCAATACATATC      587

Qy  336  AGCAGCCGCGAATGGAGTTAACAGATCTAAGCTAAAGCGATGGAATCATCTTAAAGACTTCACT      395
               |   |  ||||       | ||||  |     ||   |  ||  |||   | |||||
Db  586  ATAATCCCTAAAAAAAAAATCATTATTAAATAAAAAAAATACAATCATCTTATAAAA          527

Qy  396  ACAAGACTATGCCAAGAT-TGAAGCTAGTTATCATCTGCTTATAGTGAAGCTGAAAACAG       454
           | ||  ||     || | |  |   ||  ||| | |      |   |||  ||||||||
Db  526  AAAAATCTAAATAATAATATTAAACAAATCTAAATAAAAAACAAAAAAAATAAAAAAAA        467
```

```
Qy  455  TTAACAATAACCTTAATGCAACATTAGAACTAGAACACTAAAAATGGCTAAAAAACTAATTTAGAAT  514
           |||||||||| |||   ||||||||||   |   ||| |||||| |||| |||||||||||| |||
Db  466  AAAAATAAAACCAATAAACAATAAAAAAATAAAAAAAAAAATAACGTAATAAAACATTAATTCC      407

Qy  515  CAGCCATCAACCAAGCTAATACGATAAAAACGATAAAAAACGACTTTGATAAATGAACACCCAAATTTAG  574
              || || |  |||  |  ||| |  ||| ||   ||    | ||    || |    |    |  ||||
Db  406  AACTTACATCAATACATAAAATATCTCAATATTCTACTATAAAAAAAAAAATAAAAATTC          347

Qy  575  TTGAAGCATACAAAGCACTAAAACCACTTAGAACAACGTGCTACTAACCTTGAAGGTT        634
           |   | ||| || |  ||   | |||||| | | ||   |  ||  ||| ||  ||
Db  346  TTCAACCTAAATAATAATAAAAATCTAATTTATCATAATACTTTAAATATTATAAATC        287

Qy  635  TGTCATCAACTGCTTATATCAAATTCGCCAATAATATAAAATTCTCAATAATACTACAACTATTATCAAAGCTA  694
           |     | ||| ||||   |||  |   |||| ||||||   ||||||||||  |||||||||| |||||
Db  286  TTCAACTTCTAATTAAATAATAAATAAATAAAATTCTCAATAATACTACAACTATTATCAAA       227

Qy  695  GTAGTTAATAA  706
           ||| || |||
Db  226  TAAATATAAAAA  215
```

RESULT 10
US-10-311-455-528/c
; Sequence 528, Application US/1031145 5
; Publication No. US2003014 3606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 528
; LENGTH: 8136;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-528

```
Query Match            5.3%;  Score 68.6;  DB 15;  Length 8136;
Best Local Similarity 44.1%;  Pred. No. 0.016;
Matches  375;  Conservative   0;  Mismatches  471;  Indels   4;  Gaps   2;
```

```
Qy    2  AAAACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAAA        61
           ||||| || |  ||| ||| |||||     |||| ||||||| ||||||||| |||||
Db 1591  AAAATACTATTTTAATTAATTACTATACTACAATATAAAAACCCCAAAAATAAACTTAACAT   1532

Qy   62  ATCCTAAATAAATAAGACCGTAATTTAACTAACTAAAAATTAAAATAAATAGGTTTTTTCTTATCA     121
           |||| || |||| |   || || |  ||||| ||        || |||    || |||||  |||
Db 1531  TTAATAATAATAAAAAAATCAAAAAAATCAAAAAATATATTAAAAATTTAAAATTTTTTCCAAATAA   1472

Qy  122  ACCAAAATTCTCTAGTAATAAACCGCTTATTTTTATTTTTAGTAATCTTTTAAGATA         181
           |||||  |    ||| ||| || |   |||  |||| |  ||| |||   |||| ||
Db 1471  AAAATAATTTAACATATTAATAACATATATATATATAAAACACCCCCGTTATATAATAATTACCA    1412

Qy  182  TAAATATCTTAATATTCTATGAATAAGAAAAGAAATCATCTTAAAGACTTAGTTTGT       241
           |||| || || |  || || |||  || |||     |||| |||||| || ||  |
Db 1411  TATACAAAATAAAAAAACTACAATTAATAATATAATCTTATAAACATACATAAATATAT       1352

Qy  242  TAGGTACAACATCGTTCTTCTTAGCATTGGGATTTCTCAGCTGTGATGTCTATTACTACTAAAAAAG  301
           ||| |||| |||   |  | ||  |  |  |||| ||||||| ||||  || |  | ||| |||||
Db 1351  TAAATCAAAAAATAAAGCTATAATTTATATAAAACTACTCCTACTAAAATTAATAATAAAAAA       1292
```

```
Qy    2  AAAACATCAGATTGTTAATCTGATATCTTTGCTAAAAAAACACAAAAATCTTCTAACAAAG    61
         |||||||  | |||||     |||||||||||  |||||||| | ||||||||  ||||||
Db 1591  AAAATACTATTTTAATTATACTATAACTACAATATATAAACCCCAAAATAAACCTTTAACAT  1532

Qy   62  ATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAATTAAAAAATGTTTTCT  121
         |||  ||||| ||||  |||||||| ||||||||||||||||||||||||||||||||||| 
Db 1531  TTAATAATAATAAAAAAAAATCAAAAAATATATAAATATTAAAAAATTTTTCTCAAATAA    1472

Qy  122  ACCAAAATCTCTAGTAATAAACGCTATTATTTTTATTTTTAGTCATCTTTTAAGATA    181
         |||||||| ||||  |||| |||||||||  || ||||||||||||||||||||||
Db 1471  AAAATAATTTAACAATATTAATTATATAATAAACACCCCTCGTTATATAATTAATTACCA  1412

Qy  182  TAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTGT   241
         ||| |||||  |||||||||  ||||||||||||||| |||||||||||||||||||||
Db 1411  TATACAAAAATAAAAAACTACAATTAATATAATTAAAATCTTATAAAACATAACATAATATAT  1352

Qy  242  TAGGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTGTATGTCTTCATTACTAAAAAG   301
         |||||| || |||| |||| |||||||||||  ||||| |||||||||||||||||||||
Db 1351  TAAATCAAAAAATAACGCTATAATTATAAAACTACTCCTACTAATTAATAATAAAAAA    1292

Qy  302  ATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGCGAATGGAGTTAACAGATC   361
         ||| ||||| |||||||  ||||||||  ||||||||  |||| ||||||||||||||
Db 1291  ACAAAACAAAATTAACAACTACTAAAAATCATCACATAACTATCAAAAAAAAAATA    1232

Qy  362  TAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGAGATTGAAGCTA    421
         ||||| |||||||  | |||||||| |||  ||||||||||||||||||||||
Db 1231  TTTAAAAACCTTATATTAATTAATTAACTATAAATACTCGTTTATAATTTAAATTTA  1172

Qy  422  GTTTATCATCTGCTTATAGTTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAG   481
         |||||  | |||| || ||  || |||||||||||  |||| ||||||||||||||
Db 1171  CAAAACAAAATAATTAAAAAAATAATCTATAAAAACAATTATTATTAAAAAATTACCACTAA  1112

Qy  482  AACAACTAAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATA   541
         ||||||||||  ||||||||||| ||||  ||||| |||||||  ||||||||||||
Db 1111  TAAAACTTCAAAAAAATTTAAA-TAAAAAAAATTTAA-TAAAAAAAACTAATATAATATATA  1053

Qy  542  AAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCA   601
         |||||||||| ||||  ||||||||||||||  |||||| |||||||||||||||||||
Db 1052  AAAAAAATAATAATAATATTTTAATCAATTTAACACAAACAAAAAAACTTATTATAATTA   993

Qy  602  CTTTAGAACAACGTGCT--ACTAACCTTGGAAGGTTGTCATCAACTCTTATAATCAAA   658
         |||| |||||||||     ||||||||||||||||||||||||||||||||||||||
Db  992  TTTCTCAAAAATAATAAAAAAAAATAACTATAAACTCTATTAAACTAAAAAATTAAAAA   933

Qy  659  TTCGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTTAATAACTAAAACACTAG   718
         |||||||| | | | || |||||| ||||||| |||||||||||||||||||||||
Db  932  TTAAAATAATTTATATATATACATCCTATAACCATCTAATTTTAAAAATACTACTATAAAAACTA   873

Qy  719  ATCCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACACAGTTAATCGGAATA   778
         ||||||| | |||||||| ||||  | ||||| ||||| ||||||||||||||||||||
Db  872  TTAACCTAACAATTACGACCTCTATATAAACTTTTAAAACTTTTAAAAAATCCATAAA   813

Qy  779  TTAATAATACGTTATCAACTATTAATGAACAAAAGACTAATGCTGATGCATTATCTAARTA   838
         |||||| ||| ||||||| ||||||| |||||||| |||||||||||||||||||||
Db  812  AAAAATATTTTAATATAAATACTAATAATCATATATAAAAAACTATTAAACCAAA   753

Qy  839  GTTTTATTAA   848
         ||||||||||
Db  752  AAATTATTAA   743
```

```
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 102
; LENGTH: 11836
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (7603)
US-10-239-676-102
```

```
Query Match          5.2%;   Score 68.4;  DB 15;   Length 11836;
Best Local Similarity  44.3%;  Pred. No. 0.021;
Matches  371;  Conservative   0;  Mismatches 461;   Indels   6;  Gaps   2;

Qy      37 AAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAA 96
           |||||||| |||||||||||||   || |||| |||||| |||  ||||| ||||||
Db    4357 ATAAAAATATATATATAAACTATATAAATATATATTATATATATTTATATATATAT 4298

Qy      97 ATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTT 156
           |||||||||| ||||| |||||||| ||||||||||||| || ||||| |||  |||||
Db    4297 ATAAATTTATATTTATTATTAATTTATATTTAATAAATTTATTTTATTA 4238

Qy     157 TTATTTTAGTCATCTTTTAAGATATAAATATCTTAATATTCTATGAATAAGAAAGA 216
           |||||||||   ||||  ||||  ||||||||| ||||||||||  |||||||
Db    4237 TAAATTTTATATAAATATACTCTAATATATTTTTACATATATATATTATATACTATATAT 4178

Qy     217 ATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCT 276
           |||||||||||||||||| | |||| |||||||||||||  ||   |||||
Db    4177 ATAAAAACAAACATATAAAATAATATACATATATAT--ATATATATATATATATATATAT 4120

Qy     277 AGCTGTATGTCTATACTAAAAAAGATGCAAACCCAAATAAATGGCCAAACCCAATTAGAA 336
           ||||||||  ||| ||||||||| ||| |||||||| |||||| ||||||||||||
Db    4119 ATATATAAAATCATTTATTCAACAATATCCAAACTCTATCCTAATTACTT 4060

Qy     337 GCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTA 396
           ||||||||  ||||||||||| |||| ||    ||| |||| ||||||||||  ||||||
Db    4059 ACCATACAACAATAACAATATAAAATATAAAACCTCCCCTCTAAATAAAAAAAAAAACAA 4000

Qy     397 CAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACAGTT 456
           ||||||||||||||||||||||||||| ||||||  || ||| ||| ||| | ||||||
Db    3999 TCAAATAA-----AATATAATATATCATATAATACAAAGCCTATAAAATAATAACCAAATA 3944

Qy     457 AACAATAACCTTAATGCAACATTAGAACAACTAAAAATGCTAAAACTAATTTAGAATCA 516
           ||||||| || || |||||||||| |||| |||||    ||||||  ||||||| ||||
Db    3943 AAAAAAAAAAATTTCTAAAGAAAAATTTACTATTTTTATATAAAAAATTAAAAAAA 3884

Qy     517 GCCATCAACCAAGCTAATACGGATAAAAACGACTTTGATAATGAACACCCAAATTTAGTT 576
           ||||||||||||| ||||  | ||| |||| || |||| ||| ||||| || ||||||
Db    3883 ACTTCGATATAAAATAACATTTAAACAAAAACTTACCGAAAATAAAATAACAACCATATAA 3824

Qy     577 GAAGCGATAACAAAGCACTAAAAAACCACTTAGAACAACCACATAAAACCTAAAAAACTATA 636
           |||||  |||||||   ||||||||||||||||||| ||||| |||||||||
Db    3823 CTATATAAAAAAAAAACTAATCAAAAAAACCATAAAAACCTAAAAACCTAAAAAACTATA 3764

Qy     637 TCATCAACTGCTTATAAATCCAATTCGCAATAAATTTAGTGAACAATATAAAACAATAAAAGCTAGT 696
           ||||||||||||| ||  ||||||||| || ||| ||  |||||||   ||||| |||
Db    3763 AAATTCCTATCATATTAACAAAAAAAAACCAATATAACTTAAAAACAAAATAAATA 3704

Qy     697 AGTTTAATAACTAAAACACTAGATCCACTAAATAAAAAATTAAAAAATTAACATCTTATAAAA 756
           ||| ||| ||||  |||| ||||| || |||| |||| |||| |||||| |||||
Db    3703 AAAAAAAAATTATAAAAATTAAAAATAAATTAACAAAAATTAACATCTTATAAAAA 3644
```

```
RESULT 13
US-10-240-453-114/c
; Sequence 114, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; TITLE OF INVENTION: With DNA TRanscription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 114
; LENGTH: 11836
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (7603)
US-10-240-453-114
```

```
Query Match          5.2%;   Score 68.4;  DB 15;   Length 11836;
Best Local Similarity  44.3%;  Pred. No. 0.021;
Matches  371;  Conservative   0;  Mismatches 461;   Indels   6;  Gaps   2;

Qy      37 AAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATAAAGCCGTTAAATTAACTAAAAA 96
           |||||||| |||||||||||||   || |||| |||||| |||  ||||| ||||||
Db    4357 ATAAAAATATATATATAAACTATATAAATATATATTATATATATTTATATATATAT 4298

Qy      97 ATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTT 156
           |||||||||| ||||| |||||||| ||||||||||||| || ||||| |||  |||||
Db    4297 ATAAATTTATATTTATTATTAATTTATATTTAATAAATTTATTTTATTA 4238

Qy     157 TTATTTTAGTCATCTTTTAAGATATAAATATCTTAATATTCTATGAATAAGAAAGA 216
           |||||||||   ||||  ||||  ||||||||| ||||||||||  |||||||
Db    4237 TAAATTTTATATAAATATACTCTAATATATTTTTACATATATATATTATATACTATATAT 4178

Qy     217 ATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTCTTAGCATTGGGATTTCT 276
           |||||||||||||||||| | |||| |||||||||||||  ||   |||||
Db    4177 ATAAAAACAAACATATAAAATAATATACATATATAT--ATATATATATATATATATATAT 4120

Qy     277 AGCTGTATGTCTATACTAAAAAAGATGCAAACCCAAATAAATGGCCAAACCCAATTAGAA 336
           ||||||||  ||| ||||||||| ||| |||||||| |||||| ||||||||||||
Db    4119 ATATATAAAATCATTTATTCAACAATATCCAAACTCTATCCTAATTACTT 4060

Qy     337 GCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTA 396
           ||||||||  ||||||||||| |||| ||    ||| |||| ||||||||||  ||||||
Db    4059 ACCATACAACAATAACAATATAAAATATAAAACCTCCCCTCTAAATAAAAAAAAAAACAA 4000
```

```
Qy     757 ATTACTACAGTTAATGGAATATTAATAATACGTTATCAACTATTATGATGACAAAAGACT 816
           |||||   |||||||| ||||||||||| ||  |||| ||||||| ||||| ||||||||
Db    3643 ATTATAAAATTATACTTCTATATTAAAACTCGACTATCCAAAAAATTAAAAAAAATAAAATT 3584

Qy     817 AATGCTGATGCATTACTAATCTAATAGTTTTATTAAAAAGTGATTCAAAATAATGAACAA 874
           |||| |||||| ||||  ||||| |||   |||||||||  |||| |||||||| |||||
Db    3583 AAAAAACTATACCAATAAAATTCATTAATATATCCATTAATAAATAAACAAAAACAAAAAAAA 3526
```

```
Db   6685 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6626

Qy    156 TTTATTTTTAGTCATCTTTTAAGATATAAAATATCTTAATATTCTATGAATAAGAAAAG  215
          ||| |||||||  | |||||  |  ||| ||||||||||||||||||||| ||||||||
Db   6625 AAAAAAAAAAAAAAAAAAAAACGAAAAAAAAATAACATATATAACCATATATA       6566

Qy    216 AATCATCTTAAAGACTATTAGTTTGTTGGTAGGTACAACATCCTTCTTAGCATTGGGATTTC  275
          || ||||||| ||| |||
Db   6565 AACAAACTAAAACCCTTATACACTACTAAAAAGATGCAAACCCAAATAATGGCAACCCAATTAAA  6506

Qy    276 TAGCTGTATGTCTATTACTAAAAAGAGACTTTGATAATGAACACCCAAATTTAGTT  335
          ||||    ||| |||||
Db   6505 AAACAATATAACAATTCCTCAAAAATTAAAACATAAAAAATTACCACTATAAAACAACAATC  6446

Qy    336 AGCAGCGCGAATGGAGTTAACAGATCTCAATCAATGCTAAAGGCGATGACATTAGCTTCACT  395
          ||| ||||  | ||||
Db   6445 ACACTTCTAAAATATATACCCAAAAAAACTAAAAAAAACAAAAATTCAAACAAACATTTATACA  6386

Qy    396 ACAAGACTATGCCAAGATTGGAGCTAGTTTATCATCGTCTTATAGTGAAGCTGAAACAGT  455
          || |||||||| | |||
Db   6385 CCAATACT---CAAAATAACATAATATTCACAATAACCAAAAAACTAAAAACAACCCAAA  6330

Qy    456 TAACAATAACCTTAATGCAACATTAGAACAACATTAGACAAAATGGCTAAAAACTAATTTAGAATC  515
          ||| | ||| ||
Db   6329 TATTCATCAAC--AAATAATATACAAAATATATACAAGATATAACAATATACAATATAATATTATTC  6272

Qy    516 AGCCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAATGAACACCCAAATTTAGT  575
          ||   |  |||  |
Db   6271 AACCATAAACACTATACTATCATCAACTAAAATATATATACT--------ACAACATAAATAAACT  6222

Qy    576 TGAAGCATACAAAGCACTAAAAACCACTTTAGAACACGTGCTACTAACCTTGAAGGTTT  635
          | || || |
Db   6221 TTAAAACACTATACTATACTACCAACTACAAAAATAACTAATTATTATTATATTTCC  6162

Qy    636 GTCATCAACTGCTTATAATCAATTCGCAATAATTCGCAATAATTTAGTGGATCTATACAATAGAAAGCTAG  695
          | |||  |
Db   6161 ACTTTATATAAAATACCAAAATAAAACAAAACTAATAAAAACAAACAAAAATAAAAATAA  6102

Qy    696 TAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGA  755
          ||  | ||
Db   6101 AAAAAAATTAAAAAAAAAAAAAAACAATTATAAATTATAAAAATTCTATTTAAAAATAATAATTA  6042

Qy    756 GATTACTACAGTTAATCGGAATATTAATAATACGTTATCAACTATTAATGAACAAAAGAC  815
          ||
Db   6041 AATACTAAAAAATAATAACTTAATAAAATAAATTACACAACACTATAAATACACACCTAAAAATA  5982

Qy    816 TAATGCTGATGCATTATCCTAATAGTTTTATTAAAAAAGTGATTCAAAAATAATGAACAAA  874
          ||    ||  |
Db   5981 TCTAAATAATAAATTCCAAATTATATAATTTCACAATTAAAAAAAATAAAAAATAAATTAA  5923
```

Query Match   5.2%; Score 68.2; DB 15; Length 8392;
Best Local Similarity 45.3%; Pred. No. 0.019;
Matches 380; Conservative 0; Mismatches 443; Indels 16; Gaps 3;

```
Qy     36 AAAAAAAACACAAAATCTTCTAACAAAATCCTAAATAAAATCCTAAATAAATAAGCCGTTAAAATTAACTAAAA  95
          ||||||||  |  ||                            ||| ||||
Db   6745 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAA  6686

Qy     96 AATTAAAAAAAATGGTTTTCTTATCAACCAAAATTCTCTAGTAGTAATAAACGGCTTATTTATT  155
          || |||||                 |  |||
```

```
;    2000-06-30
;    2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 101
;   LENGTH: 12405
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (7895)
US-10-221-613-101
```

```
Query Match      5.2%;   Score 68;  DB 13;  Length 12405;
Best Local Similarity   43.3%;  Pred. No. 0.025;
Matches 365;  Conservative   0;  Mismatches 475;  Indels   2;  Gaps   1;
```

```
Qy      35 TAAAAAACACAAAATCTTCTAACAAAATCCTAATAAATAAGCCGTTAAATTAACTAAA 94
              ||||   ||  | ||||  |  ||      || ||  | ||| |||||
Db    9308 TTAACATTCTAAATATATTTTAAAATATAACAACAAATCCCTAAAAAAAATCAA 9249

Qy      95 AAATTAAAAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTAT 154
              |||        || ||   ||    |  || || ||||| |||    |
Db    9248 AAATAACTTCAATTTTAATAAAAACAATTACAAATAAAATATACTACAAAAAAAACCAA 9189

Qy     155 TTTTATTTTTAGTCATCTTTTAAGATATAAAATATCT--TAAATTCTATGAATAAGAA 212
              ||         || ||      | |||| | ||   ||
Db    9188 TTTAAAACACTAAAATAAATAAAAAAACTAACCTTAACCTTCTAACATCCACAAAAAAAAAA 9129

Qy     213 AAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTTCTTAGCATTGGGAT 272
              || ||| || || |||  || || |  | |     ||||       |||
Db    9128 CAAAACCACTAAACACGAAAATATATACTTAAAAAAAAATCCAAACTAAAAATATAAAT 9069

Qy     273 TTCTAGCTGTGATGTCTATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATT 332
              || |  ||| | ||| ||       ||   |  ||       | |  ||||
Db    9068 TCAATAATCTTCAACGAATAAACAATTTATACCACCAAAATAAAATCCACTAATTTAACAT 9009

Qy     333 AGAAGCAGCGCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTC 392
              |  |||||  |||  | |     | ||  | ||  |||||
Db    9008 AAAAAATAAACATATAAAAAAATTAAAAACCAACTCTAAAAAAATTTAAAATCCAAAA 8949

Qy     393 ACTACAAGACTATGCCAAGATTGAAGCTAGTTGTTATCATCTGCTTATAGTGAAGCTGAAAC 452
              || ||  | ||  ||   |||||  |||| |    | ||
Db    8948 AATAAAAAACAACCTACCCACTAAAAAAATCAAAAAAATATCAAAAAATCTTAAAA 8889

Qy     453 AGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGA 512
              || ||         | ||| | ||      ||  ||   ||||
Db    8888 TCCAAAAAAAAAAATATAAAAAATATAAACCATATCAAATATTACTACTTACTATCAAACAAATT 8829

Qy     513 ATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTT 572
              ||||  | |||||      || | |  |   || |  | | ||||||
Db    8828 AATAAAAAACTAATCAATAATTTACCTTTAAAAAAATCATAAAAAATRACTTTAATCAAAACTA 8769

Qy     573 AGTTGAAGCATACAAAGCACTAAAAAAACCCACTTTAGAACAACACGTGCTACTAACCTTGAAGG 632
              ||||  |  || ||  |||  |||||      ||      |  ||| ||
Db    8768 TTTTAACAAAGAACTAAAAACTAAAAAAATCTTAAATAAAAATATAAAAAATTAAAAAA 8709

Qy     633 TTTGTCATCAACTGCTTATAATCAAATTCGCAATAATTAGTGGATCCATCAATTACAATAAAGC 692
              |||||         ||||      |||  |  | ||||||
Db    8708 AAAACTAAAATAAACAACTCTTAAAAATCTTAAAAAATTCTACTACAAAAAAAAAAATAAAAC 8649

Qy     693 TAGTAGTTTAATAACTAAAACCACTAGATCCACTAAAATGGGGGAACGCTTTTAGATTCTAA 752
              ||||||||        |    | |||  ||   ||||||
Db    8648 AATATTTTTAAAAACTATAAAATCTCCAAAAATTTTTTTTTCCTTTAAAAAAATAAA 8589

Qy     753 TGAGATTACTACAGTAATCGGAATATTAATAAATACGTTATCAACTATTAATGAACAAAA 812
              || || || |  ||   |||| ||    | |     || | |
Db    8588 AACTATTTCAACATATATATATCTAATAAAATATTTAATAAAAAAAATCATT 8529

Qy     813 GACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACA 872
              ||||     | |  |||   ||  ||||| | |||||   ||   |  | |
Db    8528 AACAATACAAAAAAAAAAAAAAAAAATAACTACTAAAAAATAACAATAAAAATAAAAAAAACAAAA 8469
```

```
Qy      873 AA  874
            ||
Db     8468 TA  8467
```

```
Search completed: May  6, 2004, 00:19:34
Job time : 524.265 secs
```

```
            GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     May  5, 2004, 10:50:12 ; Search time 2985.35 Seconds
                              (without alignments)
                              13063.796 Million cell updates/sec

Title:           US-09-901-572a-1
Perfect score:   1306
Sequence:        1 aaaaacatcagattgttaat........taaatcgtttatcaggtt 1306

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        27511289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  em_gss_hum:*
                18:  em_gss_inv:*
                19:  em_gss_pln:*
                20:  em_gss_vrt:*
                21:  em_gss_fun:*
                22:  em_gss_mam:*
                23:  em_gss_mus:*
                24:  em_gss_pro:*
                25:  em_gss_rod:*
                26:  em_gss_phg:*
                27:  em_gss_vrl:*
                28:  gb_gss1:*
                29:  gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                      SUMMARIES
                                                                  %
Result                            Query                                            Match          DB       ID          Description
No.    Score                                              Length  DB                                                   -----------
-----  -----                                                                                                           ----------------
    1   79.4                                               6.1    1101    29       CNS0039G     AL063921  Drosophil
c   2   79.2                                               6.1    1200    13       BX437758     BX437758  Drosophil
c   3   78.2                                               6.0    1392    29       CG757503     P052-4-C0
c   4   78.2                                               6.0    1531    29       CG748014     P041-4-B0
```

```
c     5    6.0   78       1124   13   BX436282    BX436282 BX436282
      6    5.9   77.2      1101   29   CNS003BD    AL064091 Drosophil
      7    5.9   77.2      1200   13   BX437758    BX437758 BX437758
      8    5.9   77.2      1201   13   BX446437    BX446437 BX446437
c     9    5.8   76.8      1348   29   CG749499    CG749499 P043-4-A0
     10    5.8   76.4      1201    9   AL565455    AL565455 AL565455
     11    5.8   76.2      1101   29   CNS00EVL    AL069706 Drosophil
c    12    5.8   75.8      1101   29   CNS000B8    AL063632 Drosophil
c    13    5.8   75.6       641   28   AQ946120    AQ946120 Sheared D
     14    5.8   75.4      1225   28   CC238324    CC238324 CH261-192
     15    5.7   75        1201    9   AL536104    AL536104 AL536104
c    16    5.7   74.8      1201   29   CNS0167M    AL106396 Drosophil
c    17    5.7   74.2       998   13   BX436885    BX436885 BX436885
     18    5.7   74.2      1201   13   BX439779    BX439779 BX439779
     19    5.6   73.4      1146   29   CNS021G2    AL176843 Tetraodon
c    20    5.6   73.2      1364   29   CG757970    CG757970 P053-2-C0
     21    5.6   73        1201   13   BX420717    BX420717 BX420717
c    22    5.5   71.8      1201   13   BX356896    BX356896 BX356896
c    23    5.5   71.4      1896   29   CG753083    CG753083 P048-1-C0
c    24    5.5   71.2       960   13   BX346155    BX346155 BX346155
c    25    5.5   71.2      1566   29   CG757757    CG757757 P053-1-D0
c    26    5.5   71        1781   29   BX145762    BX145762 Danio rer
     27    5.4   71        1056   13   BX415058    BX415058 BX415058
     28    5.4   71        1091   13   BX424950    BX424950 BX424950
     29    5.4   70.6      1056   13   BX415058    BX415058 BX415058
c    30    5.4   70.2      1491   29   CG753221    CG753221 P048-2-A0
c    31    5.4   70        1373   29   CG750869    CG750869 P045-2-E1
     32    5.4   69.6      1626   14   CF238805    CF238805 AGENCOURT
     33    5.3   69.4      1092   29   CNS020K7    AL175696 Tetraodon
     34    5.3   68.8      1101   29   CNS00EO7    AL069440 Drosophil
     35    5.3   68.8      1101   29   CNS016LI    AL106896 Drosophil
c    36    5.3   68.6      1074   28   BZ696936    BZ696936 SP_Ba009
     37    5.3   68.6      1201   13   BX343343    BX343343 BX343343
     38    5.3   68.6      1201   13   BX361152    BX361152 BX361152
c    39    5.3   68.6      1592   29   CG750135    CG750135 P04-4-3-D0
     40    5.2   68.4      1187    9   AL514085    AL514085 AL514085
c    41    5.2   68.4      1210   29   CG749728    CG749728 P044-1-C0
c    42    5.2   68.4      1353   29   CG744812    CG744812 P037-3-B0
     43    5.2   68.2      1098   13   BM803974    BM803974 AGENCOURT
     44    5.2   68.2      1200   13   BX414560    BX414560 BX414560
c    45    5.2   68.2      1528   29   CG753854    CG753854 P049-1-D0
```

                                      ALIGNMENTS

```
RESULT 1
CNS0039G
LOCUS       CNS0039G      1101 bp      DNA      linear    GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION     AL063921.1   GI:4941778
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Oscegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
```

```
LOCUS       BX437758     1200 bp    mRNA    linear    EST 15-MAY-2003
DEFINITION  BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01
            5-PRIME, mRNA sequence.
ACCESSION   BX437758
VERSION     BX437758.1  GI:30773605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Peng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ InVitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CSOCAP008CA01QP1.
FEATURES             Location/Qualifiers
     source          1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSOCAP008YB01"
                     /tissue_type="THYMUS"
                     /clone_lib="Homo sapiens THYMUS"
                     /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."

ORIGIN
Query Match         6.1%;  Score 79.2;  DB 13;  Length 1200;
Best Local Similarity 30.8%;  Pred. No. 0.0024;
Matches  242;  Conservative 176;  Mismatches 343;  Indels 25;  Gaps 3;

Qy   66 TAAATAAATAAGCCGTAAAATAACTAAAAAATTAAAAAAATGGTTTTCTTATCAACCA 125
           ||  |: |:    |                 ||:  |::
Db 1043 TAATTWATHYTATTWTCTATAWTAWATAWATAAAAHWWWTTTYTWAWAAAAAAWWT 984

Qy  126 AAATTCTTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTAAGATATAAA 185
           |||       |||:||:      |||  |||  |::|    |||  ::|
Db  983 TTTTWWYATATTAAAAAAWATTWTTWTCTTTTTWTTTYTYWWTWTWAAATW 924

Qy  186 TATATCTTAAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTGTTAGG 245
           |       :||:||| ||:         |:  ::  |::|
Db  923 TTTTTATAWAAWTTTTATTTTTTTTTCMWYYTTYTTTTATTTTTTTTYTMTTATA 864

Qy  246 TACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGC 305
           |   :: |: | ::|||:| : |||:       |:|::::
Db  863 TATAAAWATATATWTWYYWTWYYYWTTTTTTTTWTWAWATMTTTYTCTWW--------- 817

Qy  306 AAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGAATGGAGTAACAGATCTAAT 365
                 |::|   |: |: :|:     :|
Db  816 -----WAATTTTTTTTAWAAWWWTTTTTMHATCCTWTWWAAAAAATAWWTATAAAAAA 763

Qy  366 CAATGCTAAAGCGGATGACATTAGCTTCACTACAGACATGCCAAGATGTGAAGCTAGTTT 425
           : |||  || ||:|  ||||:||:|  | | :|: ||:  :| |||:  ::|||
Db  762 HTCYTYCYYYATAWWTWAAWATCYYTCMATARAAAAAWTTYCTTCYCYTTTTCTWWTWT 703

Qy  426 ATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACA 485
           |:|    :| |||  ::    |||| :|||  : :|:|   |:|
Db  702 TTWAWANATYYMYTTCCTWTWTAWWAWAAWWTATMTAYYTCHYWWAAACAAWARAAAA 643

Qy  486 ACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACCGGATAAAC 545
           |:|||||  |     |||:   ||  |:  :|  | : |  |||  |||  ::
Db  642 AAWAAAAAAWAWAWMWHCYTTWAWAAAAAAAAAATCCTYMWWATAAAMWAAAWATMWT 583
```

```
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08KI0"
                     /clone_lib="RPCI-98"
                     /note="end : TE73"

ORIGIN
Query Match         6.1%;  Score 79.4;  DB 29;  Length 1101;
Best Local Similarity 19.1%;  Pred. No. 0.0023;
Matches  135;  Conservative 294;  Mismatches 272;  Indels 6;  Gaps 2;

Qy   19 ATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCTAACAAAATCCTAAATAAATAAGC 78
           |:    :|:|| ||  ||   ::|    |          |   |  |
Db  398 ATAWAWWWWWTTTTTTTTAWAAAWAAAATAAATTWAAWAWAWAAAAATTWWAAAAWAAAW 457

Qy   79 CGTTAAAATTAACTAAAAAATTAAAAAAAATGGTTTTTCTTATCAACCAAAAATTCTCTAGTA 138
            :|| : ||:|| ||                |:| ||||||  | ||   ::|  :|||||||  |
Db  458 AWTAWWTTTTAWTWAAAAAAAAAAAAAAAATTWTTTTTTTTWTTTWTTTWTTWWWTTAAAW 517

Qy  139 ATAAACGCTTATTTATTTTATTTTTAGTCATCTTTAAGATATAAATATATCTTAATAT 198
           |:        |:||: ::||: | |   |:  :  |::|
Db  518 AAAAAAAAAWAAAWAAAWATAAATWTWWTWWTYT---TWAAWATPAAAMCMAAWYYH 573

Qy  199 TCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTGTTAGGTACAACATCCTTT 258
           | : | : |  ::|:|:||:| :| |:|:|  : |::| :| |:
Db  574 TYTYTYHYTYTWYTMTYHYTVTMTWYHTAWAHTTWYHWYHTYAMWHWMTWHWTMWAWWHWT 633

Qy  259 CTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAAT 318
           ||: | |  : :|  ||  : |: | :|  |   :|| |  : |  |
Db  634 TYTAAYYYYYTCMYYHYMHWHHAHAHAAWWTHTWWTWWTHAYHWATTHYYYYMYCAMMCM 693

Qy  319 GGCCAAACCCAATTAGAAGCAGCCGAATGGAGTTAACAGATCTAAGCAGATTAACAAGCG 378
           |:| |  :|:|  ||   |    |     | : :|:|  |   || | :|   |  | : |
Db  694 CTHTCHHCYYYYHHYTAHHTHTHHWYAHYYMWYMWAYYWWYCTACYVHYHHHHHYEWAY 753

Qy  379 ATGACATTAGCTTCACTACAAGAGCATGCCAAGATGAAGCTAGTTTATCATCTGCTTAT 438
           :|   |::|   ||   | |:   |:|   ::|| | : ||  :| |:  |:|
Db  754 HTTWYAWAHAMMWHHAHYAAAAAWAA--WATTHHYHHTHYMEHTYMYHYMYTCCYMC 811

Qy  439 AGTGAAGCTGAAACAGTTAACAATAACCTTAATGCAACATTAGAACACTAAAAATGCT 498
           | : | |  |:  ||:|  |  :| | || |:| |   |  |   | | |   ||  |
Db  812 TYHCWHYHTAYTCWTWTHWMWTWTHHWMWTWYHHTWWHHTTHWAWWHTHTWCWWWHHAATTWT 871

Qy  499 AAAACTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAAACGACTTTTGATAAT 558
           | ::|||  : | || : |    | |  :|:|  | ||:  |   ||| |||  | ||| :|
Db  872 WATHCWACWTMWHWHWHWMHWMHHHHMCHAHHHTHRMCMCHHHMHCTCHHHHTMYHMTCHWMW 931

Qy  559 GAACACCCAAATTAGTTGAAGCATAACAAAGCAAGAAGATCAAAGACCACCACGTGCT 618
           | : | |: || ||:|:| |  |:|  | |   |  :|:|  | |:|  | | |||  |
Db  932 HMEHWHWHHHWWMATWMTTWTTMWMMCCMWHHCHMYHMWMEMYMYCCHYCTCHTHATTHYH 991

Qy  619 ACTAACCTTGAAGGTTTGTCATCAACTGCTTATATCAACAATTCGGCAATAATTTAGTGGAT 678
           | : | |    ||  |: |   | || |  ||| |: || |   | |||||::|||| |:|
Db  992 YMCTCYHYCWHTYWTAYWAWTAHAMTTATWWWMHWWMHWWAHWATWWWWWWATAWACT 1051

Qy  679 CTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACT 725
            | : |  :| ||  |:| |  ::| | :|   | | || |  |  |
Db  1052 CHHWTWYHTHCTWYYHHTYHMWANWMAWHWHHMYAHYHWAHHCWYY 1098

RESULT 2
BX437758/c
```

```
RESULT 3
CG757503/c
LOCUS       CG757503              1392 bp    DNA      linear   GSS  24-OCT-2003
DEFINITION  P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG757503
VERSION     CG757503.1  GI:37986131
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1392)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
  PUBMED    12884007
COMMENT     Contact: Sommer RJ
            Max-Planck-Institute for Developmental Biology
            Evolutionary Biology
            Spemannstr. 37-3, Tuebingen D-72076, Germany
            Tel: 0049707160137l
            Fax: 0049707601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1392
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

Query Match      6.0%;   Score 78.2;  DB 29;  Length 1392;
Best Local Similarity   42.8%;  Pred. No. 0.0032;
Matches   503;  Conservative   0;  Mismatches  658;  Indels   13;  Gaps    2;

Qy     36 AAAAAAACAAAAATCTTCTAACAAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAA     95
          ||||||||||  || ||||||||||||||||||||||||||||||  |||||||| ||||
Db   1392 ATAAAAAAAATAAAATAAATAAATAAATAAATAAATTAAATAATAATAATATAAATAAAATA  1333

Qy     96 AATTAAAAAATGGTTTTCTTATCCATCCAAATTCTCTAGTAATAAAACGGCTTATTATT    155
          ||||||||
Db   1332 CTTGTTTTTWTWAWWAA-AAATATAWATWWTWTARAAAAAAAWWTWTAAAAWATWAWAWTAYCYCCCW   1273

Qy    156 TTTATTTTAGTCATCTTTAAGAATAAT... (sequence)                     215
Db   1272 AANAATAATATTTATATAATAATAA...                                  1213

Qy    216 AATCATCTTAAAGACTATTAGTTTGTTAGTACTAAACACTAGATCCACT            275
Db   1212 AAAAAAAAAAAAAAATATAA--------------AATAAATATATAAAATAAAAATAAA    1165

Qy    276 TAGCTGTATGTCTATTACTAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGA   335
Db   1164 ATATTATTATTAAAATAAATATAAAAAATNAAATAAAATTAAAAAATATAAAA        1105

Qy    336 AGCAGCGCGAATGGAGTAACAGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACT   395
Db   1104 ATAAAAAAAAATATATAAAAAATTAAAATATATATTAATAAATAAAAATATTAAAT      1045

Qy    396 ACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCGTCTATATAGTTGAAGCAGT     455
Db   1044 AAATTATTAATATATATAAAAAAATAAAATGGCTAAAACTAATTTATATATAAATC       985

Qy    456 TAACAATGAACCTTAATGCAACAT...                                   515
Db    984 AAAATATTAAAAAAAAT...                                          925

Qy    516 AGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAAATGAACACCCAAATTAGT   575
Db    924 TATAAAAAANAAAATAATTATAAAATTAATAAAATAAATAATAAATATAAATATATATTAT   865

Qy    576 TGAAGCATACAAAGCACTAAAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTT   635
Db    864 AAAAAAAAAAAAATTAAAAATATATTATTAAATAATAAATAATAATAAAATAATAATA     805

Qy    636 GTCATCAACTGCTTATATATCAAATTCCGCAATAATTTAGTGGATCTATACAATAAAGCTAG   695
Db    804 AAAATAAAAAAAAATAAAAATTAAAATATATATAATATAATTAATTATTAAATAAATAAATA   745

Qy    696 TAGTTTAATAACTAAAAACACTAGATCCACTAAATGGGGAACGCTTTTAGATTCTAATGA   755
Db    744 AAAAAAAAAATAATAAAAAAAAAATATAAAAAAAATATAAAAAAAAATATAAAAAAT      685

Qy    756 GATTACTACAGTTAATCGGAAATATTAATAATACGTTATCCAATAATTAATGAACAAAGAC   815
Db    684 TATTAAATAAAAATATAAAATIIIIIAAAATATATATTATATAATTTAAATAAAATAAAA   625

Qy    816 TAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGAACAAAAG   875
Db    624 ATATTTATAAAATAAAAAAATNTATAATAAAATAATAATAAAATATAAAAAATATAAAAAT   565

Qy    876 TTTTGTAGGGACTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTTGTTGCTTT    935
Db    564 TTTTAAATAAATTTATAATTTATTTATTAATTATTAATTATATATATTATTTAT        505

Qy    936 TAGTGCTGATGTAACACCCGTCAATTATAAAATGCTAGAAGGACCGTTTGGAATGGTGA    995
Db    504 ATAATATATAAAAAAAAAAAAAAATAATTAAAAAAAAAAATAATTAAAAAAAAAATA      445

Qy    996 TGAACCTTCAAGTAGAATTCTTGCAAACACGAATAGTATCACAGATGTTTCTTGGATTA   1055
Db    444 ATAATAAAAAAAAAAAAAAATAAAT-AAAATAAAAATAAAATTAAAATATAATATATTA    386

Qy   1056 TAGTTTAGTCGGAACAAACACGAAGTACCAATTTAGTTTTTTAGCAACTATGGTCCATCAAC  1115
Db    385 AAATAAAAAAAATAATAAAAAATAAAATATAATTAAATAAAAAAAAAAATAATAATTAA    326

Qy   1116 TGGTTTATTTTATATTCCCTTATAAGTTGGTTAAAGCAGCTAATAACGTTGGATT       1175
Db    325 TTTATATTTATTATTATTAAAATATATAATTTATAAAATTTAAAATTTATATA         266

Db   1176 ACAATACAAATTAAAATAAATAAATTGAAAATGTTCAACAA                     1209
```

Db   265 AATATATAAAAAAATATTAAATAAATTAAAATAA 232

RESULT 4
CG748014
LOCUS   CG748014   1531 bp   DNA   linear   GSS 24-OCT-2003
DEFINITION P041-4-B02.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG748014
VERSION   CG748014.1 GI:37968940
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1531)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
  COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 0049707160171
            Fax: 0049707601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source 1..1531
             /organism="Pristionchus pacificus"
             /mol_type="genomic DNA"
             /strain="California"
             /db_xref="taxon:54126"
             /clone_lib="Ppa EcoRI BAC Library"
             /note="The library was generated by a partial digest of
             the genomic DNA with EcoRI and cloning into the BAC
             vector."
ORIGIN

Query Match   6.0%;   Score 78.2;   DB 29;   Length 1531;
Best Local Similarity   39.4%;   Pred. No. 0.0031;
Matches 344;   Conservative 0;   Mismatches 530;   Indels 0;   Gaps   0;

Qy   1 AAAAACATCAGATTGTTAACTGATATCTTTGCTTAAAAAAACACAAAATCTTCTCAACAA 60
Db   489 AAAAANAAAANAAAANAAAAAAAAAAAAAAAAAAAANAAAAAANAAANAAAAAAAAAAAA 548
Qy   61 AATCCTAAATAAATAAATAAGCCGGTTAAATTAACTAAAAAAAATTAAAAAAATGGTTTTTCTTATC 120
Db   549 AANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNA 608
Qy   121 AACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTAGTCATCTTTTAGAT 180
Db   609 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAA 668
Qy   181 ATAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTTTG 240
Db   669 AAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAA 728
Qy   241 TTAGGTACAACATCCTTTCTTAGCATTGGGATTCTCAGCTGTATGTCTATTACTAAAAAA 300
Db   729 ANAAAAAAAAAAAAAAAAAAAAAAAAANNAAANAAAAAAAANNAAAAANNAAAAAAAA 788
Qy   301 GATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCCGCGAATGGAGTTAACAGAT 360
Db   789 AAAAAAAAAAAAAAAAAAAAANAANAAAAAAAAAAAAANAAAAAAAAAAAAAAANAAAAAA 848
Qy   361 CTTAATCAATGCTAAAGCGGATGACATTAGCTTCACTACAAGACTAATGCCAAGAGATTGAAGCT 420

Db   849 AAAAAAAANAAAAAAAAAAANAAAAAAAAAAANANANANANAAAAAAAAAN 908
Qy   421 AGTTTATCATCTGCTTATAGTGTAAGCTGAAGCAGTTAACAATAACCTTAATGCAACATTA 480
Db   909 AAAAAAAAAAANNNAAAAAAAANNAAAAAAAAAAAAAAAAAAAAAAAAAAAA 968
Qy   481 GAACAACTAAAAATGGCTAAACTAATTAGAATCAGCATCCATCAACCAAGCTAATACCGAT 540
Db   969 NAAAAAAAAAAAAAAAAAAATAAATAAAAATAAAAAANNAAAANAAAANANANAAAAAAAAA 1028
Qy   541 AAAACGACTTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACC 600
Db   1029 AAAAAAAAAAAAAAAAAAAAAAANAAANANANAAAAAAAAAAANNAAAAAAAAAAAAAN 1088
Qy   601 ACTTTAGACACGTGCTACTAACCTTGAAGGTTGTCATCAACTGCTTATATCAAATT 660
Db   1089 NAAANANNAAAAANNAAAAAAAAAANANAANANAAAAAAAAAAAAAAAAAAAAAAAA 1148
Qy   661 CGCAATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTAATAACTAAAACACTAGAT 720
Db   1149 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAAAAAAAAAAAAAAAAAN 1208
Qy   721 CCACTAAATGGGGGAACGCTTTAGATTCTATAATGAGATTACTACAGTTAATCGGAATATT 780
Db   1209 AANNAANAAAAAAANAAANAAAAAAAAAAANANAAAAAAAAAAAAAAAAAAAAAAAA 1268
Qy   781 AATAATACGTTATCAACTATTAATGACAAAAGACTAATGCTCGATGCATTATCTTAATAGT 840
Db   1269 AAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAANNAAAAAAAAAAAAAAAAAAA 1328
Qy   841 TTTATTAAAAAGTGATTCAAAATAATGAACAAA 874
Db   1329 NAAANAAAAAAAAAANAAAAAAAAAAAAAAAAAAANA 1362

RESULT 5
BX436282/c
LOCUS   BX436282   1124 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YC01
            5-PRIME, mRNA sequence.
ACCESSION   BX436282
VERSION   BX436282.1 GI:30787521
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1124)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fullength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP001AB01QP1.
FEATURES   Location/Qualifiers
     source 1..1124
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0CAP001YC01"
             /tissue_type="THYMUS"
             /clone_lib="Homo sapiens THYMUS"
             /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
             with a NotI-oligo(dT) primer. Five prime end enriched,
             double-strand cDNA was digested with Not I and cloned into
             the Not I and EcoRV sites of the pCMVSPORT 6 vector.
             Library was not normalized."

Left column (RESULT 6)

```
ORIGIN
Query Match      6.0%;  Score 78;  DB 13;  Length 1124;
Best Local Similarity   24.2%;  Pred. No. 0.0038;
Matches 203;  Conservative 255;  Mismatches 371;  Indels 10;  Gaps 2;

Qy    83 AAATTAACTAAAAATAAAAAATGGTTTTCTTATCAACCAAAATTCTCTAGTAATAA 142
         ||  |  ::|||| |||||||  |: ||||  | |||||||:||||||:||
Db  1122 AAYYYYYYAAAAAAAAAAAAHHHYYYYYYHAAAAAAAAAAHHAAAHYYHAYAAA 1063

Qy   143 ACGCTTATTTATTTTATTTTAGTCATCTTTAAGATCATAAATATCTTAATATTCTA 202
         |: ::|: ||:  ::::|:|  | :: ::::||  :  ::::::|:|:||:
Db  1062 AYNYYYYYHAAAYYYAYYYYHAYYYYYYYYYYYYYYYHAAAAYYYY 1003

Qy   203 TGAATAAGAAAAGAATCATCTTAAAGACTATAGTTTGTTAGGTACAACATCCTTTCTTA 262
         |:  :||||| |||  ::  ::: :::  ::||||||||YYYYYYHAA
Db  1002 YYYYYYAAAAAAAAAAAAHHMAAYYYYMYHYHYYYYYYAAAAAAAAAAYYYYYYHAA 943

Qy   263 GCATTGGGGATTTCTAGCTGTATGCTATTACTAAAAAGAGATGCAAACCCAAATAATGGCC 322
         :  ::::: ::  : |||| |: ||||||:|||||||HAAAAAAAHYYYY
Db   942 :AYYYYYYYYYYYYYYAYYYYHAAAAAHHHHHAAAAAAHAAAAAAAHYYYY 883

Qy   323 AAACCCAATTAGAAGCAGCGGCGAATGGAGTTAACAGATCCTAATCAATGCTAAAGCGATGA 382
         ||:|| :||  |||||||  :|: ||||  :   ::||||||YYYYYYYAAAA
Db   882 YYYYHHHAAAAAAYYYYHHAAAAAAAAAAAAAAYYYYYYYYYAAAA 823

Qy   383 CATTAGCTTCACTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTG 442
         :  :::|||| ||| ||: ||||||| |: ||||HAAAYYYYAAAYYYYYYYAAAY
Db   822 AAAAAAYYYYY------YYYYYVAAAAAAAHAAAAAAYYYYAAAYYYYYYYAAAY 769

Qy   443 AAGTCGAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAA 502
         |||||: ||||||  ||||||||:  |:|||||HYHAAAAAAA
Db   768 YYYAAAMYHAAAAAAAAAAHYYHAAAAAAAAAAAAHYHAAAAAAA 709

Qy   503 CTAATTTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGATAATGAAC 562
         ::: :||  :||||||||||||||||| |||:  NHHAAAAAAAAAA-
Db   708 AAAHYYYYYYYYYYYYYYYYYYYYYYHAAAAAAAAANHHAAAAAAAAAA- 650

Qy   563 ACCCAAATTAGTTGAAGCATACAAAGCACTACAAGACTAAAAACCACTTTAGACAACAGTGCTACTA 622
         |::  |||||||:||: ||||: |:|||||||YYYYHAHYYYYYYYYYAA
Db   649 --AAAHYYYYYYYYYHAHHAAAAAAAAAAAAYYYYYHAHYYYYYYYYYAA 593

Qy   623 ACCTTGAAGGTTTGTCATCAACTGCTTATAATCAACTGCCAATAATTTAGTGGATCTAT 682
         ||:||: ||||||  :||||||||||||||||HYYYYYYY
Db   592 AAAHYHAMYYYYYYHAAAAAAAAAAAAAAAAHYYYYYYY 533

Qy   683 ACAATAAAGCTAGTAGTTTAATAACTAAAACTAGATCCACTGAAATGGGGGAACGCTTT 742
         :|||||HAAAAAAAAAAAHHAAAAAAAAAAVYAAAHYYHAAHHAAAAYYYYY
Db   532 HAAAAAAAAAAAHHAAAAAAAAAAVYAAAHYYHAAHHAAAAYYYYY 473

Qy   743 TAGATTCTAATGAGATTACTACAGTTAATCGGAATATTAATACGTTATCAACTATTA 802
         |||| |||||||||  :||||||  ||  |::  :|| |||YYYYYYAAAAAAAHYYY
Db   472 YYYYAAAAAAAAAAAAYYHAAAAAAHHHAAAAAAYYYYYYAAAAAAAHYYY 413

Qy   803 ATGAACAAAAGACTAATGCTGATGCATTATCTAATAGTTTATTAAAAAGTGATTCAAA 862
         |: ||||||| ||||||: || || ::|: :||| ||:THAAAAAAAAAAAAAAH
Db   412 YYYYYYYYYHAAAAAAHHHAAAAAAAYYYYYYTHAAAAAAAAAAAAAAH 353

Qy   863 ATAATGAACAAAGTTTGTAGGGACTTTTACAAACGCTAAGTTCAACCTTCAAACTAC 921
         :  ::  :::|||||:|: :||||  ::|:| ||||  ||YYYYCCCYYYYTWAAAAC
Db   352 HYYYHHHYYYYRAAHYYYYHYYYYYYYYYAAAYYYYCCCYYYYTWAAAAC 294
```

RESULT 6
CNS003BD
LOCUS   CNS003BD   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION   Drosophila melanogaster genome survey sequence TET3 end of BAC #
             BACR08K08 of RPCI-98 library from Drosophila melanogaster (fruit
             fly), genomic survey sequence.
ACCESSION   AL064091
VERSION   AL064091.1   GI:4941847
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)

Right column

```
ORGANISM     Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
             Ephydroidea; Drosophilidae; Drosophila.
             1 (bases 1 to 1101)
REFERENCE
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
             collaboration with the Berkeley Drosophila Genome Project (BDGP).
             The BDGP is constructing a physical map of the Drosophila
             melanogaster genome using these BACs. For further information
             please see http://www.fruitfly.org The BDGP Drosophila
             melanogaster BAC library was prepared by Kazutoyo Osoegawa and
             Aaron Mammoser in Pieter de Jong's laboratory in the Department of
             Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
             NY. The library is named RPCI-98 and was constructed by partial
             EcoRI digestion of Drosophila DNA provided by the BDGP from the
             isogenic strain y2; cn bw sp, the same strain used for the BDGP's
             P1 and EST libraries. A more detailed description of the library
             and how to order individual BAC clones, the entire library, or
             filters for hybridization from the BACPAC Resource Center can be
             found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES     Location/Qualifiers
source       1..1101
             /organism="Drosophila melanogaster"
             /mol_type="genomic DNA"
             /db_xref="taxon:7227"
             /clone="BACR08K08"
             /clone_lib="RPCI-98"
             /note="end : TET3"
ORIGIN
```

```
Query Match      5.9%;  Score 77.2;  DB 29;  Length 1101;
Best Local Similarity   40.4%;  Pred. No. 0.0052;
Matches 157;  Conservative 47;  Mismatches 185;  Indels 0;  Gaps 0;

Qy    13 TTGTTAATCTGATATCTTGCTTAAAAAACACAAAATCTCTAACAAAATCCTAAATAA 72
         ||| ::|: :||||||||||||| |||||:|||||||||:|:|:|:||:|||
Db   664 TTATTTAWNAAWATTAAAAAWATTAAAAAAAAATATAAAAATTATAAAWTW 723

Qy    73 ATAAGCCGTTAAATTAACTAAAAAATTAAAAAAATAAGTGTTTTTCTTATCAACCAAAATTCT 132
         |||||||| |||||||||| ||||||||:||||||||YYYYYYTTATTTATTAAATAT
Db   724 TAAATATATAWTWTTTAWAAAATATATTTAAAAAAAATATAAAAAATAT 783

Qy   133 CTAGTATAAACGCTTATTTATTTTAGTCTTTAAGTATATAATATATTN 192
         ::|||||:||||: ||:::|| ||:|:|||||||||:|:|:||::|:
Db   784 WATWATTTATATANWNWATAWTTTWTTAAATTWATAAATWTWATTAWWATTN 843

Qy   193 TAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATAGTTGTTAGGTACAACA 252
         |||| :|||| ||||| :| |||| :|  :|||||||:||||||TTGTTAGGTACAACA
Db   844 TAWAATAAAAAAWAAAAATATAAAATAATTTWATAWTTTWWAAATAAWTTTA 903

Qy   253 TCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTACTAAAAAGATGCAAACCCA 312
         |::|||::| |||||||| |||:|| ||| ||:TGTATGCTATTACTACTAAAAAGATGCAAACCCA
Db   904 TAAWAAAAAAAAATTAAWATAAAAWTWAAWTAWTTTWTTMTYTAWAAAWTTATTAWAWAW 963

Qy   313 AAAGCCGATGACATTAGCTTCACTCACAAGA 372
         ||||| |||| ::|||| ||:|:| |||||
Db   964 AATAATTTTATATWATWATTATAATAWWAAAAAATAAAATWAAATWWAWAWWAYAA 1023

Qy   373 AAAGCCGATGACATTAGCTTCACTACACAAGA 401
         ||||||||||||||||||||:||:|:|||||
Db  1024 AWAYACWAAAAAWTWTATWYATAWMAAAA 1052
```

RESULT 7
BX437758
LOCUS   BX437758 Homo sapiens   1200 bp   mRNA   linear   EST 15-MAY-2003
DEFINITION   BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01
             5-PRIME, mRNA sequence.

```
ACCESSION    BX437758
VERSION      BX437758.1  GI:30773605
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 1200)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.
FEATURES     Location/Qualifiers
     source  1..1200
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0CAP008YB01"
             /tissue_type="THYMUS"
             /clone_lib="Homo sapiens THYMUS"
             /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
             with a NotI-oligo(dT) primer. Five prime end enriched,
             double-strand cDNA was digested with Not I and cloned into
             the Not I and EcoRV sites of the pCMVSPORT 6 vector.
             Library was not normalized."
ORIGIN

Query Match           5.9%;  Score 77.2;  DB 13;  Length 1200;
Best Local Similarity 30.6%;  Pred. No. 0.005;
Matches 256;  Conservative 156;  Mismatches 418;  Indels 6;  Gaps 1;

Qy   16 TTAATCTGATATCTTTGCTTAAAAAAACACAAAATCTTCAACAAAATCCTAAATAAATA 75
Db  365 TWTWTTTTATWWWWNDDDDAKCAWADAGAAITWTTTTTTWRAGRGARKRDGTTDTRAADWW 424

Qy   76 AGCCGTTAAATTAACTAAAAAAATAAATAAAATGTTTTTCTTATCAACCAAAATTCTCTA 135
Db  425 WWTWAAWAWWWDKRKTTTTTTTTTKKRRAAAATTTTTTTTTTWTWAKAAAAARAAATW 484

Qy  136 GTAATAAACGCTTAATTTATTTATTTTAGTCATCTTTTAAGATATAAATATATCTTAA 195
Db  485 TTWWRAAAAAAAWWTWWWAARGRARGATTTTTTTAAWWWGGGAGRTAWWTWTTWTTTTT 544

Qy  196 TATTCTATGAATAAGAAAGAAATCATCTTAAAGACTATTAGTTGTTAGGTACAACATCC 255
Db  545 TTTTTTTTAWWWATWWATWTATATTTTWTWWTWAWAAAASAGAWKATWTTTKKTTTA------ 598

Qy  256 TTTCTTAGCATTGGGATTTCTAGCTGTATGCTATTACTAAAAAAGATGCAAACCCAAAT 315
Db  599 TWWKRAGGAITWTTTTTTTWTWAARGDWKWTWTWTTTTTTTTTTTTTWTTGTTTWWW 658

Qy  316 AATGGCCAAACCCAATTAGAAGCAGCGGCGGATGAGTTAACAGATCTAATCAATGCTAA 375
Db  659 RDGARRTAKATAKWKWTTWTWWTAWAWAGGAARKRRATWTWTWARAAWWAGAAAARGAGA 718

Qy  376 GCGATGACAATAGCTTCACTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCT 435
Db  719 AGRAAWTTTTTTTATKAGARRGATWTTAWWTWATRARGARGADTTTTTTATAWWTATTT 778

Qy  436 TATAGTGAAGCTGAAACGTTAAACAGTTAACAATAACCTTAATGCAACATTAGAACAACAACTAAAAATG 495
Db  779 TTTWAWAAGATDKAAAAAWWWTTWTAAAAAATTWWWAGARAAAKATWTWAAWAAAA 838

Qy  496 GCTAAAACTAATTAGAATCAGCCATCAACCAAGCTAATACGGATAAAACGACTTTTGAT 555
Db  839 AAAAAAWRARWAAATATATWTTTATATAAKARAAAAAATAAAATAAAARARRWKGAAAA 898

Qy  556 AATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACGT 615
Db  899 ARAAATARAAAAWWTTWTATAAAAAWWTTW3AWWRAAARAWAWAAAAAAAAGARAAGAW 958

Qy  616 GCTACTAACCTTGGAAGGTTGTCATCAACTGCTTAPATATCAAATTCGCAAPAATTAGTG 675
Db  959 AAWAATWTTTTTAATATRWRWAAAAAWWTTTTTTWTWARAAAAWWDTTTTATWTWTATWT 1018

Qy  676 GATCTATACAAPAAGCTAGTAGTTTAATAACTAAAACACTAGATCCACTAAATGGGGGA 735
Db 1019 AWTATAGAWAAATARDATWRAATTATWADARAAAAGAWATAKDAGAGATAAAWTTTTTT 1078

Qy  736 ACGCTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATTAATAATACGTTATCA 795
Db 1079 TTTTTTTTWTAAAAAAGAAAATWTTARAAAAAAAAAAWATAAAAWATAAAAAAAAAAA 1138

Qy  796 ACTATTAATGAACAAAGACTAATGCTGATGCATTACTAAGTTTTATTAAAAA 851
Db 1139 AAGAAKADDAAWAAGAAAATAAAAAAAAAAGWRAARAWWAARAAAAAAWTAARA 1194


RESULT 8
BX446437
LOCUS        BX446437    1201 bp    mRNA    linear    EST    22-MAY-2003
DEFINITION   BX446437 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA006ZG10
             3-PRIME, mRNA sequence.
ACCESSION    BX446437
VERSION      BX446437.1  GI:31025727
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. This sequence belongs to sequence cluster 7563.f For
             more information about this cluster, see
             http://www.genoscope.cns.fr/
             cgi-bin/cluster.cgi?seq=CLOBA006ZG10FP1&cluster=7563.f. Contact :
             Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CLOBA006ZG10FP1.
FEATURES     Location/Qualifiers
     source  1..1201
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CLOBA006ZG10"
             /tissue_type="PLACENTA"
             /clone_lib="Homo sapiens PLACENTA"
             /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
             with a NotI-oligo(dT) primer. Five prime end enriched,
             double-strand cDNA was digested with Not I and cloned into
             the Not I and EcoRV sites of the pCMVSPORT 6 vector.
             Library was not normalized."
ORIGIN

Query Match           5.9%;  Score 77.2;  DB 13;  Length 1201;
Best Local Similarity 35.9%;  Pred. No. 0.0049;
Matches 265;  Conservative 81;  Mismatches 393;  Indels 0;  Gaps 0;

Qy  147 TTATTTATTTTATTTTTTAGTCATCTTTTAAGATATAAATATATCTTAATATTCTATGAA 206
Db  101 TTTTTTTTTTTTTTTTTTTTTTTTTWWWTWTWTWTTWWWTTTTWAWWAAWAAAA 160
```

```
                                                    Tel:  00497071601371
                                                    Fax:  00497071601498
                                                    Email: ralf.sommer@tuebingen.mpg.de
                                                    Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..1348
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."

ORIGIN
        Query Match          5.9%;   Score 76.8;  DB 29;   Length 1348;
        Best Local Similarity  44.7%;  Pred. No. 0.0054;
        Matches 376;   Conservative   0;  Mismatches 457;   Indels   8;  Gaps   2;

Qy     34  TTAAAAAACACAAAATCCTCTAACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAA     93
           ||||||| ||| ||||||||||||||||||||||||||||||||||||||
Db   1254  TTATAATTATTATTTATAAATAAAAAAAAAAAATATAAAAATTAATTATATAAAAATAT   1195

Qy     94  AAAATTAAAAAAATGGTTTTCTTATCAACAACAGTGCTACTAACCT               153
           ||||| |||||             ||| ||||||||||||||||||||
Db   1194  ATAAAATAAAAAAANAAAAAAAAATATTNTATAAAATATAATATATTTTAAATATATTTTA   1135

Qy    154  TTTTTATTTTAGTTCATCTTTAAGATATAAATTCATCTAATTCTATGATAATAAGAAA    213
           ||| |||||||||||||||||||           ||||||
Db   1134  TATAATAATAAAAAATATATAAAAAAAATAATATTATTAAAAAAATATAAAAAAAATTAA   1075

Qy    214  AGAATCATCTTAAAGACTATTAGTTTGTTGTTAGGTCACAACATCCTTCTTAGCATTGGGATT    273
           ||||              |||||||||||||||
Db   1074  TAAATAATAAAAAAATTTAATTTNATATTTTATATATTTTTATAAAATAAAATAAAAAA   1015

Qy    274  TCTAGCTGTATGTCTATTACTAAACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTA    333
           ||||||||||||||||||||  |||                |||
Db   1014  TTAAAATAATATTTAAAAATTNATTAAAAATTATTAATTATTWNAAARTAATTTAAT    955

Qy    334  GAAGCAGCGCGGAATGGAGTTAACAGATCCAATCAATGCCGATGACATTAGCTTCA    393
           ||||||||||||||||||||||||    :::|:|||||||||||||
Db    954  TAAAAAAAAATAATAAAAATAATAAAAATAATAAAAAAAAAAAAAAAAAA    895

Qy    394  CTACAAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTGAAGCTGAAACA    453
           |::|| |::||                     |||||||||
Db    894  AATAAAAAAAATATAANAATAAAAATATATTATATATAAAAAATAA---AAAA    839

Qy    454  GTTAACAATAACCTTAACTAAATGCAACATTAGAACAACTAAAAAATGGCTAAAAACTAATTTAGAA    513
           ||||||||||||:||||||||            |||||||||
Db    838  AAAAAAAAAAAATTTAATTAATAAAAAAAATAATAAAAAAATTTATATAAAAATAATAAAAAATA    779

Qy    514  TCAGCCATCAACCAAGCTAATACGGTAATACGGATAAAACGACTTTTGATAATGAACACCCAAATTTA    573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778  AAAAAAAATAAAAATTATTATAAAAAAAATATATTTTTTAAAAATTAAANAAAAAATAATAA    719

Qy    574  GTTGAAGCATACAAAGCACTAAAAACCACTTTAGAACAACAGTGCTACTAACCCTTGAAGGT    633
           ||||||||||||||||||
Db    718  AAAAAAAATATATATATAAAAAAAAATATATATAAAAATAAAAAAAATAAAAAA---    662

Qy    634  TTGTCATCAACTGCTTATAATCAAATCGGAATATCAAATTTAGTGTGGTCTACACAATAAAGCT    693
           |||            ||||||
Db    661  -TAATAATAATTAAAAAAATAATATATTAAAAAATATTATATAAAAAAAT    603

Qy    694  AGTAGTTTAATAACTAAAACACTAGATCCCACTAAATGGGGGGAACCGCTTTAGATTCTAAT    753
           |||         |||||
Db    602  ATAAATAAAAAAATATTAATTAATAAAATAAAAATAATTNTAAATATTNTAAAAAAAA    543

Qy    754  GAGATTACTACAGTTAAATCGGAATATTAAATACGTTAATACAACTATTAATAATGAACAAAAG    813
           |||            ||||| |
Db    542  ATAAATAATAAAACTAAAATAATAATATAAAATATATAAAAAATAAAAAAAAAAAAAG    483

Qy    814  ACTAATGCTGATGCATTATTCTAAAGTTTTATTAAAAAAGTGATTCAAAATAAATGAACAA    873
```

```
Qy    297 AAAAGATGCAAACCCAATAATGGCCAACCCAATTAGAAGCAGCGGCGAATGGAGTTAAC 356
              |: :||  | ||||: ||||   :||:|: ||   |||||||||||||
Db    670 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 729

Qy    357 AGATCTAATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCAAGCATTGGA 416
              ||::|| ||   :|| |: ||||: | :|||||   |:|||
Db    730 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 789

Qy    417 AGCTAGTTTATCATCTGCTTATAGTGTGAAGCTGAAACAGTGAACAATAACCTTAATGCAAC 476
              |   |||:| ||| ||: :|: | |  | | :|  : :|
Db    790 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 849

Qy    477 ATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAGCCATCAACCAAGCTAATAC 536
              : ||||||| ||  ||: | | | |||  |::|   ||   ||||  || |||
Db    850 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 909

Qy    537 GGATAAAACGACTTTGATAATGAACACCCAAATTTAGTTGAAGCATACAAAGCACTAA 596
              |  |||||| | | |||  || | |   |   ::|||
Db    910 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 969

Qy    597 AACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATCAACTGCTTATATATCA 656
              | |||| | | ||||   |   |||  |  |: || ||
Db    970 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 1029

Qy    657 AATTCGCAATAATTTAGTGGATCCATACAATAAAAGCTAGTAGTTTAATAACTAAAACACT 716
              ||| ||| ||   | |||| || |||  |  ||||
Db   1030 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 1089

Qy    717 AGATCCACTAAATGGGGAGACGCGTTTTAGATTCTAATGAGATTACTACAGTTAATCGGAA 776
              | |||| |||   | :||   | ||| |||| | ||||
Db   1090 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 1149

Qy    777 TATTAATAACGTTATCAACTAATTAAATGAACAAAAGACTAA 818
              : ||||| | |  ||:: |||  :|| : |||
Db   1150 ATAWATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 1191


RESULT 11
CNS00EVL
LOCUS       CNS00EVL      1101 bp    DNA      linear    GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION     AL069706.1 GI:4949849
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
```

```
Db    482 TAAAATAAAAAAAAAAATTAAAAATAAAAAAAAAATATAAATAATTAAAAAAAAAATTAAAAAA 423
              |||      |||    |||             || :|||   |||| || |||||  |||| ||
Qy    874 A 874

Db    422 A 422


RESULT 10
AL565455
LOCUS       AL565455      1201 bp    mRNA      linear    EST 12-MAY-2003
DEFINITION  AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF005YO18 3-PRIME, mRNA sequence.
ACCESSION   AL565455
VERSION     AL565455.2 GI:30549492
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12916848.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9232.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.
FEATURES             Location/Qualifiers
source       1..1201
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="CS0DF005YO18"
             /tissue_type="FETAL BRAIN"
             /dev_stage="fetal"
             /clone_lib="Homo sapiens FETAL BRAIN"
             /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
             was primed with a NotI-oligo(dT) primer. Five prime end
             enriched, double-strand cDNA was digested with Not I and
             cloned into the Not I and EcoRV sites of the pCMVSPORT 6
             vector. Library was not normalized."
ORIGIN
Query Match    5.8%;   Score 76.4;   DB 9;   Length 1201;
Best Local Similarity  31.8%;   Pred. No. 0.0066;
Matches  242;   Conservative 122;   Mismatches 398;   Indels   0;   Gaps   0;

Qy     57 ACAAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAAATAAAAATCGTTTTTCT 116
              |:||   ||||||::||  || :||||  |:| | |:  |:|||||||  ||  |:||
Db    430 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 489

Qy    117 TATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTA 176
              |:||   ||||||  :||:|: || ||  |:|  |  |::|||||  |  |:|:
Db    490 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 549

Qy    177 AGATATAAATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAG 236
              | :||||  ||:| | || ::|  |  ||| |  ||||
Db    550 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 609

Qy    237 TTTGTTAGGTACAACATCCTTTCTTAGCATTGGGATTTCTAGCTGTATGTCTATTACTAA 296
              :: ||||| | |||| |  |: |:|   |:||  || ||||   |::| |:| || ||
Db    610 ATAWATAWATAWATAWATAWATATAWATAWATAWATAWATATAWATAWATATAWATATAWATATAW 669
```

AUTHORS   Genoscope.
TITLE   Direct Submission
JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
           - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
           collaboration with the Berkeley Drosophila Genome Project (BDGP).
           The BDGP is constructing a physical map of the Drosophila
           melanogaster genome using these BACs. For further information
           please see http://www.fruitfly.org The BDGP Drosophila
           melanogaster BAC library was prepared by Kazutoyo Osoegawa and
           Aaron Mammoser in Pieter de Jong's laboratory in the Department of
           Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
           NY. The library is named RPCI-98 and was constructed by partial
           EcoRI digestion of Drosophila DNA provided by the BDGP from the
           isogenic strain y2; cn bw sp, the same strain used for the BDGP's
           P1 and EST libraries. A more detailed description of the library
           and how to order individual BAC clones, the entire library, or
           filters for hybridization from the BACPAC Resource Center can be
           found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR01A24"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN

Query Match   5.8%;   Score 75.8;   DB 29;   Length 1101;
Best Local Similarity 38.4%; Pred. No. 0.0086;
Matches 198; Conservative 67; Mismatches 249; Indels 1; Gaps 1;

Qy   59 AAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATAAAAAATGGTTTTTCTTA 118
          ||||| :| || |||||||||| : |||
Db   1092 AAAAAAWAWAAAATTAAATATTTWWAAAAAAAAAAATATAAAWATATWWTWWAWTA 1033

Qy   119 TCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTATTTTTAGTCATCTTTTAAG 178
          :|:| ||| : : |||||||| ||:||
Db   1032 ATAAAAAATAAAATTWTATATAAAATWWTAAWATAATATWTTTAATAAAWTTAAAW 973

Qy   179 ATATAAAATATATCTTAATATTCTATGAATAAGAAAAGAATCATCTTAAAGACTATTAGTT 238
          ||:|: :||:| | ||||||| | |||
Db   972 ATAWATTWAWAAAAAAAAAAAAAAAAAATAAAATAAAWTTWTAAWATWAAAAWAT 913

Qy   239 TGTTAGGTACAACATCCTTCTTAGCATTGGGATTCTAGCTGTATGTCTATTACTAAAA 298
          || : | ||| :|| |||||
Db   912 ATAATTTAAWWATWAAAAAWTTAAWTWWTAAATTTAATTTTTTTTTTAWTAAAA 853

Qy   299 AAGATGCAAACCCAAATAAATGGCCAAACCCAATTAGAAGCAGCGGCGGAATGGAGTTAACAG 358
          ||| | | :| |||||||
Db   852 AAAAAAWAAAAAAATAAAATWAAAAWWWATAAWWWAAWATAAWAAATWAAAAAA 793

Qy   359 ATCTA-ATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACATGCCAAGATTGAA 417
          | | ||: || | :|: |||:
Db   792 HTAAATWAAAATTACMAAAAATTWAATTYAAAATAAATWWWATAAWTWTWAWAWWATAWAW 733

Qy   418 GCTAGTTTATCATCTGCTTATAGTAGTGGAAGCTGAAGCTAGAAACAGTTAACAATAACCTTAATGCAACA 477
          | |||: | ||:| | ||| |||||||||
Db   732 AAAAAATTAWATTHCMAAATAAWCAAAAAAWCAAAAAWAWCAMAAAAAAAATATTTTTYCCAAWA 673

Qy   478 TTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCACCAAGCTAATACG 537
          ||: ||:|:| |||||||| |
Db   672 TCACAAYMACCAACACAAAATCAMACHAAACACCAAYYAAWCAMAAYCACAAAATYBCAAAT 613

Qy   538 GATAAAACGACTTTTGATAATGAACACCCAAATT 572
          | ||:|:| | |||||||||
Db   612 CAYYAAAGRSDCYYAAAYAAATKSYHHAWATTTT 578

RESULT 13
AQ946120/c

---

     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29923"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

Query Match   5.8%;   Score 76.2;   DB 29;   Length 1101;
Best Local Similarity 32.0%; Pred. No. 0.0074;
Matches 209; Conservative 121; Mismatches 321; Indels 2; Gaps 1;

Qy   42 ACACAAAATCTCTAACAACAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAATTAA 101
          :::::::| :| :: :: :::::::: ::::: ||:|| :||
Db   443 MMMMMMHAMATYYCTCAHTWTMMMMMMWWAATWTWAAWAAAWTTATWAATWAAAAAAWA 502

Qy   102 AAAAATGGTTTTTCTTATCAACCAAAATTCTCTAGTAATAAACGCTTATTTATTTTTATT 161
          :::|||| |||| :|: :|: | :|:||| ||||| :: |||||| ||
Db   503 WWWATTTTTWWWWTWAYTWTTTWAWWTWTWAWWTWTWAWTTAAAAAWATAAAAWWAATA 562

Qy   162 TTTAGTCATCTTTTAAGATAATAAAATAATATCTTAATATTCTATGAATAAGAAAAGAATCAT 221
          : |:|| || :|||| |:| :|||| | ||||
Db   563 WATTAAWAATTTAWAAWTATATTAATWTATTAATWTWATAATAAAAAAATATTTTT 622

Qy   222 CTTAAAGACTATTAGTTTGTTAGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTG 281
          : ||||| : ||| :: ||| ||||
Db   623 TWATAAAATTTTTAATAATTTAATTAWTTAAATAAWTWATWWWTAAATTAAAAATAATT 682

Qy   282 TATGTCTATTACTAAAAAAGATGCAAACCCAAATAAATGGCCAAAATAAATGGCCAAATAGAAGCAGC 341
          |: | ||||||| :|| ||| | || ||| |:| |||
Db   683 TWAAATAWAAAAAAAAAAAAAAAAAWATWAAWAATWATAWATAWAWAAAWAATTAAAAAWAATAA--AAW 740

Qy   342 GCGAATGGAGTTAACAGATCTAATCAATGCTAAAGCGGATGACATTAGCTTCACTACAAGA 401
          ||| ::| :|: |||: :| ::|
Db   741 AAWAAATWAWWWATAATATWWTATATATWTTAWWWATWTTATAWWWTATAWAAATAWAAW 800

Qy   402 CTATGCCAAGATTGAAGCTAGTTCATCTGCTTATCATCTGCTTATAGTGAAGCTGAAGCTGAAAACAGTTAACAA 461
          :|: | |:| |:| ||| || ||||
Db   801 AWWATAAATATWWTATWAWAWAAWAWAWAWWTTWTTWTWAWWTATATWTATWWWAAAAAATWTAA 860

Qy   462 TAACCTTAATGCAACATTAGAACAACTAAAAATGGCTAAAAACTAATTTAGAATCAGCCAT 521
          ||: |:   :|: : ::|||||||:|| |:|||| ||||
Db   861 TATWAAWTATAAWAAAAAAAAATAWWTTWTTTWWAWWTATATAAAWATAWAAAAWAAA 920

Qy   522 CAACCAAGCTAAATACGGATAAAAACGACTTTTTGATAATGAACACCCAAATTTAGTTGAAGC 581
          ||| :|||: |||| ||||
Db   921 AAAAAAAAATAAAWAWWWTWATATWTTTATTAAAWTWTATTTWWAWTWTATAW 980

Qy   582 ATACAAAGCACTAAAACCACTTTAGAACAACGTGCTACTAACCTTGAAGGTTTGTCATC 641
          |||::|| |||||| |||
Db   981 TTWTAWTTATATWTAWTAWATAWATATTTTATTTAAWWTATATTTTAWAAAWTAATATATW 1040

Qy   642 AACTGCTTATAATCAAATTCGCAATAAATTAGTGGATCCTATACAATAAAGCTA 694
          ||:||:   :| :   ::| |||
Db   1041 ATAWWTAWATATAWWWAAATTAWTWTATATATWTAATAWWAAAAAAWAAAWA 1093

RESULT 12
CNS000B8/c
LOCUS   CNS000B8   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR01A24 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063632
VERSION   AL063632.1 GI:4938680
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)

```
LOCUS         AQ946120             641 bp    DNA      linear   GSS 27-JAN-2000
DEFINITION    Sheared DNA-46J23.TR Sheared DNA Trypanosoma brucei genomic clone
              Sheared DNA-46J23, genomic survey sequence.
ACCESSION     AQ946120
VERSION       AQ946120.1  GI:6769385
KEYWORDS      GSS.
SOURCE        Trypanosoma brucei
  ORGANISM    Trypanosoma brucei
              Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
              Trypanosoma.
REFERENCE     1  (bases 1 to 641)
  AUTHORS     El-Sayed,N., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
              Gerrard,C., Leech,V., de Jong,P., Uilu,E., Melville,S.,
              Donelson,J., Fraser,C. and Adams,M.
  TITLE       Determination of clone end sequences from Trypanosoma brucei GUTat
              10.1 sheared DNA library
  JOURNAL     Unpublished (1999)
COMMENT       Other_GSSs: Sheared DNA-46J23.TF
              Contact: Najib M. El-Sayed
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: nelsayed@tigr.org
              Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
              DNA library constructed at TIGR. Clones will be available for
              distribution through ATCC. Sheared DNA end sequences search page:
              http://www.tigr.org/tdb/mdb/tbdb/.
              Seq primer: M13-Reverse
              Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..641
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927/4 GUTat 10.1"
                     /db_xref="taxon:5691"
                     /clone="Sheared DNA-46J23"
                     /clone_lib="Sheared DNA"
                     /note="Vector: pUC18; Site 1: SmaI; Constructed at The
              Institute for Genomic Research (TIGR), Rockville, MD.
              Genomic DNA isolated from a cloned population of
              Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
              sheared to give a tight size distribution (approx 2 kb).
              The v + i method used for library construction is
              described in detail in Smith, H.O. and Venter, J.C.
              (Making small insert libraries for whole genome shotgun
              sequencing projects. In Genome Sequencing: A Practical
              Approach, eds. M. Vaudin and B. Barell, Oxford University
              Press, 1999)."
ORIGIN

Query Match         5.8%;  Score 75.6;  DB 28;  Length 641;
Best Local Similarity   46.0%;   Pred. No. 0.012;
Matches  255;   Conservative   0;   Mismatches  299;   Indels   0;   Gaps   0;

Qy    750 TAATGAGATTACTACAGTTAATCGGAATATTAATATACGTTATCAACDTTATGAACA 809
              ||||||||| |||  || |||  ||||
Db    624 TAGTAATAGTACTACTAATAATAATAATAATAATAATAATAATAATAATAATAATA 565

Qy    810 AAAGACTAATGCTGATGCATTATCTAATAGTTTTATTAAAAAAGTGATTCAAAATAATGA 869
                ||  |||||||   |||| ||||
Db    564 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 505

Qy    870 ACAAAGTTTGTGTAGGGACTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTTTGT 929
                |||||||    |||| ||||
Db    504 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 445

Qy    930 TGCTTTTAGTGCTGATGTAACACCCGTCAATTATAAATATGCAAGAAGGACCGTTTGGAA 989
               |||  ||||    |||| ||||
Db    444 TAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 385

Qy    990 TGGTGATGAACCTTCAAGTAGAATTCTTGCAAACACGAAGTATCCACGATGTTTCTTG 1049
                ||||| ||||||| ||||    || |||   |||| ||||
Db    259 AATTTTAAAAATAAAAATANNAANNNNNNNNNNNNNNNAAAAANNNAAAATANNTNAATTNATT 318
```

```
RESULT 14
CC238324
LOCUS         CC238324       CH261-19212_RM1.1 CH261       1225 bp    DNA        linear   GSS 12-MAY-2003
DEFINITION    CH261-19212_RM1.1 CH261 Gallus gallus genomic clone CH261-19212,
              genomic survey sequence.
ACCESSION     CC238324
VERSION       CC238324.1  GI:30564987
KEYWORDS      GSS.
SOURCE        Gallus gallus (chicken)
  ORGANISM    Gallus gallus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE     1  (bases 1 to 1225)
  AUTHORS     Krenitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
              Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE       Gallus gallus BAC End Reads
  JOURNAL     Unpublished (2003)
COMMENT       Contact: Richard K. Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 182000   Std Error: 0.00
              Seq primer: RM1 TACGACTCACTATAGGGAGA
              Class: BAC ends
              High quality sequence start: 44
              High quality sequence stop: 100.
FEATURES             Location/Qualifiers
     source          1..1225
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-19212"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2; EcoRI;
              CH261 Female chicken library - for library and clone
              ordering information: http://www.chori.org/bacpac"
ORIGIN

Query Match         5.8%;  Score 75.4;  DB 28;  Length 1225;
Best Local Similarity   38.8%;   Pred. No. 0.0094;
Matches  330;   Conservative   0;   Mismatches  512;   Indels   9;   Gaps   2;

Qy      4 AACATCAGATTGTTAATCTGATATCTTGCTTAAAAAAACACAAAATCTTCTAACAAAAT 63
              | ||||| || || |||||   |||||| ||| ||||   ||||
Db    259 AATTTTAAAAATAAAAATANNAANNNNNNNNNNNNNNNAAAAANNNAAAATANNTNAATTNATT 318
```

```
Db    384 TAATATATAATAATAATAATAATATATAATAATAATAATAATAATAATAATAATAATAATAA 325
              |||   ||| |||| ||| |||
Qy   1050 GATTTATAGTTTAGCTGGAACAAACGAGTACCAATTAGTTTAGCAACTATGGTCC 1109
              |||| ||||| ||
Db    324 TAATATATAATAATAATAATAATATATAATAATAATAATAATAATAATAATAATAA 265

Qy   1110 ATCAACTGGTTATTTATTATTCCCTATAGTTGGTTAAAGCAGCAGTCAATAACGT 1169
                |||||||||   ||   |||
Db    264 TAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 205

Qy   1170 TGGATTACAATACAAATTAAATAATGAAATGTTCAACAAGTGAGTTGGCCACTTCAAC 1229
                |||| ||||   |||| ||||
Db    204 TAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 145

Qy   1230 TAGTGCAAATATACTACAGCTAATCCAACTCCAGCAGTGATGAGATTAAAGTGCTAA 1289
              |||| ||||||| |||| ||||||
Db    144 TAATAATAATAATAATAATAATAATAATAATGATGATGATAATAATAATAATAG 85

Qy   1290 AATCGTTTTATCAG 1303
              || |||
Db     84 AAGAGTGTTTGAG 71
```

```
JOURNAL      Unpublished (2001)
COMMENT      On Feb 13, 2001 this sequence version replaced gi:12799597.
             Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Library was constructed by Life Technologies, a division of
             Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
             http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
             Faraday Avenue Genoscope sequence ID : CS0DF022BB09QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF022YC18"
                     /tissue_type="FETAL BRAIN"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /dev_stage="fetal"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
Query Match         5.7%;  Score 75;  DB 9;  Length 1201;
Best Local Similarity  34.5%;  Pred. No. 0.011;
Matches 183;  Conservative 95;  Mismatches 250;  Indels 3;  Gaps 1;

Qy    49 ATCTTCTAACAAATCCTAAATAAATAAGCCGTTAAATTAACTAAAAAAATTAAAAAAATG 108
         :|     |    | |||||   :::|||   |||:||    |||| || :::|:
Db  1184 ATATATAAYMACTAWWTRAATATWAAAATWWTHTAWAAWATWTRATAATATCAAWWWWTA 1125

Qy   109 GTTTTTCTTATCAACCAAAATTCTCTAGTAATAAAACGCTTATTTTATTTTTAGTC 168
         :|:||                :::|:|   ||||   :|| :: |||| |  ::|||
Db  1124 TWTHTATTTTTTTTTTTWWTAYWAAAATAWTWTWTTTTAAWAYWWTWWTTTAAWA 1065

Qy   169 ATCTTTTAAGATATAAAATATATCTTAAATATTCTCATGAATAAGAAAAGAATCATCTTAAAG 228
         ||::|:   ||   |||||||||:|:   ||| ::|   |||   ||:::||
Db  1064 WTATWATTAAAAAAWAWATWTTTTATWATASWATWWAAAAWAAAAWAATWATAAW 1005

Qy   229 ACTATTAGTTGTTAGGTACGAACATCCTTTCTTAGCATTGGGATTCTCAGCTGTATGTCT 288
         :|||||            |:||:|   ||:|     | |:
Db  1004 WATATTTWWAATTWTWTTTTAWATAWATWTTWTAWWWTATTTAYWCWTAWTA-TATATTW 945

Qy   289 ATTACTAAAAAAGATGCAAACCCAAATAATGGCCAAACCCAATTAGAAGCAGCGGCGAATG 348
         ||     |   ||||||   :|||| |||
Db   944 TWWTWWTAWCTAATWATWWWWAWAATWWTWWTAWAAAWTTWTATAAATAAAAWATAAAAWTT 885

Qy   349 GAGTTAACAGATCTAATCAATGCTAAAGCCGATGACATTAGCTTCACTACAAGACTATGCC 408
         |  |   ||||          |||||| :  :|||  :::||
Db   884 AAAATTAAAAWAATAAATTTTATWAAAAATAAWWAATATATWTTTWTAWTAWAATAWA 825

Qy   409 AAGATTGAAGCTAGTTTATCATCTCGCTTATAGTGAAGCTGAAACAGTTAACAATAACCTT 468
         :|     | :::  |  :  ::|:    :| |: |:|  |:||
Db   824 AAAWWTTAWWATTATWATTAATMWAAWAWAAWAWAAWATWATWWWTATAAAWTATAA----A 768

Qy   469 AATGCAACATTAGAACAACTAAAAATGGCTAAAACTAATTTAGAATCAGCCATCAACCAA 528
         |||    |         |||:||||| | | || :|||| ||:||
Db   767 AATTTAAAATATATAAAWAWTATATWAWAAWATAAWAAAAWWAAAAAWAATWATTAAWWTWATAAT 708

Qy   529 GCTAATACGGATAAAACGACTTTGATAATGAACACCCAAATTAGTTGAA 579
         ||||||   :|||||  :|   :::||
Db   707 ATAWAWAAAAAATAAAAWAWATAAWAAAAAAATTTATATAHAHTTATAAWHAA 657

Search completed: May 5, 2004, 19:58:19
Job time : 2997.35 secs
```

```
                                                                    123
Qy   64 CCTAAATAAATAAGCCGTTAAATAACTAAAAAATTAAAAAAATGGTTTTTCTTATCAAC
        |||        |||
Db  319 ATTAANNAAAATNTTNTTAAATAAAATTAAAAA--AATANNNNNNTTATTAAATANAAA  376

Qy  124 CAAAATTCTCTAGTAATAAACGCTTATTTTATTTTTAGTCATCTTTAAGAGATATA   183
        ||||  |   ||||           | :|:  |||
Db  377 ATAAANNNNAAAAAAANTANATATTTTTTTTTTNTTNNATTTNATNTANTATTTAAATA  436

Qy  184 AATATATCTTAATATTCTATGAATAAGAAAGAATCATCTTAAAGACTATTAGTTTGTTA  243
        |::  |          :   ||||||   ::    |:|:
Db  437 ATTATNNTTTANNTTTTTTTATTTTTTTAATATTAAATTAATTNAAATTTATTNTATAAAA  496

Qy  244 GGTACAACATCCTTCTTAGCATTGGGATTTCTAGCTGTATGTCTTATTACTCAAAAAAGAT  303
        ||| |  | ||  ||||
Db  497 ATAAAAAAAATNTTTTTNNAANNNNNATAAANTNNNTAAAATNNNTNNNTNANAAAAAAAAA  556

Qy  304 GCAAACCCAAATAATGCCAAATCCCAATTAGAAGCAGCCGCGAATGAGTTAACAGATCTCA  363
        ||||   |        :  ||  ||||
Db  557 AAAAATAATAANAANTTAANATNTNNAAANNTTATTATAAATAAAAANNTNNAAAAANNTNAT  616

Qy  364 ATCAATGCTAAAGCGATGACATTAGCTTCACTACAAGACTATGCCCAAGATTGAAGCTAGT  423
        |||   |  |||  |||||:
Db  617 TAAAAATTATATANNAAAATTTTNTNTANNTTATATTATTATTAAAAATAATAAAAATATTTA  676

Qy  424 TTATCATCTCGCTTATAGTGTGAAGCTGAAACAGTTAACAATAACCTTTAATGCAACATTAGAA  483
        ||||||  |  |||  |||
Db  677 NTATTNTAAATATAAAAAATNTTAATTATTTTTNAAATATATTAAAAATATAAAANTAATATAA  736

Qy  484 CAACTAAAAATGGCTAAAACTAATTTAGAATCCAGCCATCAACCAAGCTAATACGGATAAA  543
        |||||
Db  737 NATATNAAAAAAAAAATNNATTNTNATNNTATAATAATAATAATATATATATNTTTTATTAT  796

Qy  544 ACGACTTTGATAATGAACACCCCAAATTAGTTGAAGCATCATACACAAAGCACTAAAAACCACT  603
        |||||
Db  797 AAAAATAATAAATTTNNTNNNNNNNAAAATATAATATATATATAAAAANAAATTNNNNNN  856

Qy  604 TTAGACAACAAGTGCTACTACCTTGAAGGTTTGTCATCAACTGCTTATAATCAAATTCGC  663
        |||     |       || |     |||
Db  857 TNNAAANTTATTAATAAANTATTTTNTAAAATATAATTTAAATNNTNNTAAAATTNNNTNATAT  916

Qy  664 AATAATTAGTGGATCTATACAATAAAGCTAGTAGTTTAACTAAAACACTAGATCCA  723
        ||| ||                 |:|   ||||
Db  917 AATAAATTAAAANATTTAAAAAAAATTTNNATNATNNNAATNATNAAAAAAAAANNTA  976

Qy  724 CTAAATGGGGGAACGCGCTTTAGAATTCCAAATGAGATTACTAACAGTTAACAAAAGACTAATATTAAT  783
        |||||                 || | |||||
Db  977 TTAA------TATANTAATAANANNNAATAAAAATATATAAAATAAATNTATATTAATTTTAATTAAATA  1029

Qy  784 AATACGTTATCAACTATTAATGAACACAAAAGACTAAATCGCTGATGCATTATCCTAATAGTTTT  843
        ||    |  || | |||   |||
Db 1030 AAATAATATAAAANANAATAAAAATATAAAANTATTTTATAAAAAAAATTATNTAAAAATNTTTN  1089

                                                                    854
Qy  844 ATTAAAAAGT
         |||| |||
Db 1090 TATAAATAAAT  1100
```

RESULT 15
AL536104/c
LOCUS   AL536104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF022YC18 5-PRIME, mRNA sequence.
            AL536104   1201 bp   mRNA   linear   EST 31-MAY-2003
DEFINITION AL536104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF022YC18 5-PRIME, mRNA sequence.
ACCESSION   AL536104
VERSION   AL536104.2 GI:31260974
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: May 5, 2004, 09:55:26 ; Search time 11945.1 Seconds
(without alignments)
11571.357 Million cell updates/sec

Title: US-09-901-572A-2
Perfect score: 3189
Sequence: 1 atgaatatttctaaaaact.........caaccgctcccaagaataa 3189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699s residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3189	100.0	3189	6	AX665165	AX665165 Sequence
2	3145.8	98.6	3189	1	AB023292	AB023292 Mycoplasm
3	3144.2	98.6	301042	1	AB016967	AB016967 Mycoplasm
4	3143.4	98.6	8354	1	AF214004	AF214004 Mycoplasm
5	3143.4	98.6	10651	6	AX113685	AX113685 Sequence
6	3057.8	95.9	3129	6	AX665242	AX665242 Sequence
7	1052.4	33.0	7141	1	AF083976	AF083976 Mycoplasm
8	993.2	31.1	1131	1	AB033210	AB033210 Mycoplasm
9	924.8	29.0	1128	1	AB033211	AB033211 Mycoplasm
10	124.6	3.9	8760	1	MYCMGP	M31431 M.genitaliu
11	124.6	3.9	15787	1	U39698	U39698 Mycoplasma
12	124.6	3.9	110000	6	AR300198_2	Continuation (3 of
13	102.2	3.2	9691	1	MYCATTP	M21519 M.pneumonia
14	102.2	3.2	16876	1	AB000002	AB000002 Mycoplasm
15	93	2.9	1618	1	MGU34970	U34970 Mycoplasma
16	80.8	2.5	105682	3	AC116957_3	Continuation (4 of
17	80	2.5	116807	3	BX890614	BX890614 Danio rer
18	78.4	2.5	4695	1	MGU34842	U34842 Mycoplasma
19	76.2	2.4	146275	2	AL935272	AL935272 Danio rer
20	75.2	2.4	155204	2	AC007926	AC007926 Trypanoso
21	75	2.4	213544	2	BX510939	BX510939 Danio rer
22	73.6	2.3	151341	5	AL929536	AL929536 Zebrafish
23	72.4	2.3	94534	5	AL929250	AL929250 Zebrafish
24	71.8	2.3	64707	3	AC115607	AC115607 Dictyoste
25	71.8	2.3	225581	2	BX537105	BX537105 Danio rer
26	71.4	2.3	198743	2	BX530070	BX530070 Danio rer
27	70.8	2.2	3502	1	AY212515	AY212515 Mycoplasm
28	70.2	2.2	183357	2	BX569779	BX569779 Danio rer
29	69.2	2.2	146415	5	AL929469	AL929469 Zebrafish
30	68.8	2.2	266544	3	AC116956	AC116956 Dictyoste
31	68.6	2.2	151802	3	AC114263	AC114263 Dictyoste
32	68.4	2.2	182676	3	BX547933	BX547933 Danio rer
33	68	2.1	10593	3	AY232271	AY232271 Dictyoste
34	67.6	2.1	82139	3	AC115684	AC115684 Dictyoste
35	67.4	2.1	12463	3	AC115578	AC115578 Dictyoste
36	67.4	2.1	167913	2	BX537273	BX537273 Danio rer
37	66.6	2.1	111882	3	AC115612	AC115612 Dictyoste
38	66.4	2.1	110000	3	AC116984_3	Continuation (4 of
39	66.4	2.1	178664	2	BX511115	BX511115 Danio rer
40	66	2.1	176258	2	AL929105	AL929105 Danio rer
41	65.8	2.1	250029	3	AB014820	AB014820 Plasmodiu
42	65.4	2.1	169478	2	AC139280	AC139280 Homo sapi
43	65.4	2.1	195532	2	BX004760	BX004760 Danio rer
44	65.4	2.1	250195	3	AB014831	AB014831 Plasmodiu
45	65.2	2.0	157033	2	BX323881	BX323881 Danio rer

ALIGNMENTS

RESULT 1
AX665165
LOCUS   AX665165   3189 bp   DNA   linear   PAT 26-MAR-2003
DEFINITION Sequence 2 from Patent EP1275716.
ACCESSION   AX665165
VERSION   AX665165.1 GI:29290295
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum
  ORGANISM Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1
  AUTHORS   Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
  TITLE   Modified dna molecule, recombinant containing the same thing, and
            uses thereof
  JOURNAL   Patent: EP 1275716-A 2 15-JAN-2003;

```
FEATURES          Zeon Corporation (JP)
source            Location/Qualifiers
                  1..3189
                  /organism="Mycoplasma gallisepticum"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:2096"
                  /note="mgc3 gene"
ORIGIN

Query Match       100.0%;  Score 3189;  DB 6;  Length 3189;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches 3189;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy      1 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60

Qy     61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120

Qy    121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGTCTTGGACAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGTCTTGGACAA 180

Qy    181 AATGGTTCGTTGTTCAATACAGTTCCTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AATGGTTCGTTGTTCAATACAGTTCCTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240

Qy    241 AATGGAACAATTATCAAAATTAGATAGTAGTTTTACTAAAACCATTATATGTTTAGATCTAAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AATGGAACAATTATCAAAATTAGATAGTAGTTTTACTAAAACCATTATATGTTTAGATCTAAGT 300

Qy    301 GATGATTGTGGTGGATACAAAGTAAAACAAAATAGTTTCAGATTACAACTAGCAGAAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GATGATTGTGGTGGATACAAAGTAAAACAAAATAGTTTCAGATTACAACTAGCAGAAAT 360

Qy    361 AGATTTGATCAAGACAAGAGCAGAGCATATTATGCTCTGTTGGTTAATGAGAGCTAAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGATTTGATCAAGACAAGAGCAGAGCATATTATGCTCTGTTGGTTAATGAGAGCTAAC 420

Qy    421 GTTCATTTAAAAGAATTAATACTAACTAACTAACTTGGTAATAGAAACAACAATTCT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTTCATTTAAAAGAATTAATACTAACTAACTAACTTGGTAATAGAAACAACAATTCT 480

Qy    481 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540

Qy    541 ACTAAATTAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGAACTTATACATTCAAAGAAAG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ACTAAATTAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGAACTTATACATTCAAAGAAAG 600

Qy    601 GCGCCAATCTTACCTAAAGATTCACCCCAGATTGGTATAACTTATACATTCAAAGAAAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCGCCAATCTTACCTAAAGATTCACCCCAGATTGGTATAACTTATACATTCAAAGAAAG 660

Qy    661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTGAGTAGTGGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTGAGTAGTGGA 720

Qy    721 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGGAATGGTCAAATAACTAATAACAGATCCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 ACAAATGCTGATGATGGGATGTTTGATTGTGTGGGAATGGTCAAATAACTAATAACAGATCCT 780

Qy    781 ATTGCTCAAACTAAAACCACTACTGATTAATCAAAATCCTTCAACTTTAATTCAGGAGCA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ATTGCTCAAACTAAAACCACTACTGATTAATCAAAATCCTTCAACTTTAATTCAGGAGCA 840

Qy    841 ATGCCTGGTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAACA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATGCCTGGTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAACA 900

Qy    901 TCTTTCCAATTAGATGATGAAAAATTGTTTTATCCGAATGGACTGGTTCTGAAGAATAAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TCTTTCCAATTAGATGATGAAAAATTGTTTTATCCGAATGGACTGGTTCTGAAGAATAAA 960

Qy    961 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 AATATTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC 1020

Qy   1021 ATACCCGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTATTTCAAGACTA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATACCCGGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTATTTCAAGACTA 1080

Qy   1081 TACTTAAACTCAGTAATTCTTTATCATTCATTGGTGATAGTATTATATTTTGGTACC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTATATTTTGGTACC 1140

Qy   1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAGATATCTGATCTAACCGCT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAGATATCTGATCTAACCGCT 1200

Qy   1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACACA 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACACA 1260

Qy   1261 AACGGAACAACGACAACAGCAGCAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGCAGCT 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AACGGAACAACGACAACAGCAGCAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGCAGCT 1320

Qy   1321 ACTACTAACACTTCTCGAAACAGTTCTCAATCCTACTTTAAATACTTAAATACTTAGT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ACTACTAACACTTCTCGAAACAGTTCTCAATCCTACTTTAAATACTTAAATACTTAGT 1380

Qy   1381 ATTGATAGTAACAAACCAACTTCTGCAAACAAATAGATGAAACTAATTGGGCAGATCCTAAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ATTGATAGTAACAAACCAACTTCTGCAAACAAATAGATGAAACTAATTGGGCAGATCCTAAC 1440

Qy   1441 GTTATTGAAGCAAGAATATATGCTCGTCTTGTTGGTTAATGATGAAGCTAAC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GTTATTGAAGCAAGAATATATGCTCGTCTTGTTGGTTAATGATGAAGCTAAC 1500

Qy   1501 ACTAATGCAGGAAACTTTATCCGAAACAACAATTGGTGTGTTGGTTTTACTTGGAAACTTCAA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ACTAATGCAGGAAACTTTATCCGAAACAACAATTGGTGTGTTGGTTTTACTTGGAAACTTCAA 1560

Qy   1561 TCAAGAGTAGTTTTAAGAGCTTCAACTGGATGCAACGTCCAACTGGAAACTTCCAA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TCAAGAGTAGTTTTAAGAGCTTCAACTGGATGCAACGTCCAACTGGAAACTTCCAA 1620

Qy   1621 CCTTTCTTATACGTATTTGGTTATTGGTAAACAACAGCGCCCTTACGACGTATTAGATTCTTCCAAGAGTA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCTTTCTTATACGTATTTGGTTATTGGTAAACAACAGCGCCCTTACGACGTATTAGATTCTTCCAAGAGTA 1680

Qy   1681 TACGGAACATATAAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTTCCAAGAGTA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACGGAACATATAAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTTCCAAGAGTA 1740

Qy   1741 GGTACTGAAACCAATCAATTAGAGAACTTCATTAACATCACCCTGTTATGGGTGGATAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GGTACTGAAACCAATCAATTAGAGAACTTCATTAACATCACCCTGTTATGGGTGGATAT 1800

Qy   1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATAAGAGCACAAGGTGAC 1860

Qy   1861 ACACCAGAAAGCCGAAGCATCTCCAATCGGCCAATCTGGCTATTATGAGTAGTACATT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACACCAGAAAGCCGAAGCATCTCCAATCGGCCAATCTGGCTATTATGAGTAGTACATT 1920

Qy   1921 CAATCAGTTTAGGATTTGATGGAATTGAAATAACTAACTTTAAATGTTGGGGTTAAAGCATCA 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CAATCAGTTTAGGATTTGATGGAATTGAAATAACTAACTTTAAATGTTGGGGTTAAAGCATCA 1980

Qy   1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACACA 2040
```

```
Db   1981  AGCTTCTAAACTCAAATAGACCAAATCCAACACGGTCTAGAAATGATTGCTGCAACAACA  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041  TACTTAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAAACCAACCATTC     2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  TACTTAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAAACCAACCATTC     2100

Qy   2101  GGAACAACTCACCAGTTATTCAGTATCGATCACCTGGTGATCAGTTCTCATCAATTAAGAAT  2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GGAACAACTCACCAGTTATTCAGTATCGATCACCTGGTGATCAGTTCTCATCAATTAAGAAT  2160

Qy   2161  ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT  2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT  2220

Qy   2221  AAATCTAGTGTTTATACATTAAGAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  AAATCTAGTGTTTATACATTAAGAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2280

Qy   2281  TTCAGTCCAACAAGTTAAATTGACGTTAATGAAAATTGGTGTAATCTTACCTTATTAGAC  2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  TTCAGTCCAACAAGTTAAATTGACGTTAATGAAAATTGGTGTAATCTTACCTTATTAGAC  2340

Qy   2341  AATTCATTCTATACAGTAAAATGCTGCTGCGTGGTAATGTTGCATTGTTCTCATCAAACCCTGGT  2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  AATTCATTCTATACAGTAAAATGCTGCTGCGTGGTAATGTTGCATTGTTCTCATCAAACCCTGGT  2400

Qy   2401  TCTCCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT  2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  TCTCCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT  2460

Qy   2461  GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT  2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461  GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT  2520

Qy   2521  GAGTACTTAATTCAAAATGGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA  2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521  GAGTACTTAATTCAAAATGGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA  2580

Qy   2581  AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA  2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581  AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA  2640

Qy   2641  GTTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCAAACTCGAGATCCAA         2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641  GTTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCAAACTCGAGATCCAA         2700

Qy   2701  TACTTAGATGGTAAGTATTATGATGGTAAATTAAAGAACAATAATTTAGTAACATTCTCT  2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701  TACTTAGATGGTAAGTATTATGATGGTAAATTAAAGAACAATAATTTAGTAACATTCTCT  2760

Qy   2761  TATAAACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA  2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761  TATAAACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA  2820

Qy   2821  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTAAGAGCTCAA  2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTAAGAGCTCAA  2880

Qy   2881  AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT  2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881  AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT  2940

Qy   2941  GCTGTCTTAGGTGTTCAGTTTACAAGAGATTATTACCCAAACGTTAAGAAAAACCT     3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941  GCTGTCTTAGGTGTTCAGTTTACAAGAGATTATTACCCAAACGTTAAGAAAAACCT     3000

Qy   3001  GCTGCTTTAGGTGCTCGTGTAAATCTGGTGATAAGAAACCTGCTGCTCGTGCTAAACCTGCT  3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001  GCTGCTTTAGGTGCTCGTGTAAATCTGGTGATAAGAAACCTGCTGCTCGTGCTAAACCTGCT  3060

Qy   3061  GCTCCAGCTAAACCATCTGACCAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA    3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061  GCTCCAGCTAAACCATCTGACCAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA    3120

Qy   3121  TCTGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC   3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121  TCTGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC   3180

Qy   3181  AAAGAATAA  3189
           |||||||||
Db   3181  AAAGAATAA  3189
```

```
RESULT 2
AB023292
LOCUS       AB023292                3189 bp    DNA     linear   BCT 27-MAY-2000
DEFINITION  Mycoplasma gallisepticum gene for 120kDa membrane protein, complete
            cds.
ACCESSION   AB023292
VERSION     AB023292.1  GI:4240363
KEYWORDS    120kDa membrane protein.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (sites)
  AUTHORS   Yoshida,S., Fujisawa,A., Tsuzaki,Y. and Saitoh,S.
  TITLE     Identification and expression of a Mycoplasma gallisepticum surface
            antigen recognized by a monoclonal antibody capable of inhibiting
            both growth and metabolism
  JOURNAL   Infect. Immun. 68 (6), 3186-3192 (2000)
  MEDLINE   20278096
   PUBMED   10816462
REFERENCE   2  (bases 1 to 3189)
  AUTHORS   Yoshida,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1999) Shigeto Yoshida, Jichi Medical School,
            Department of Medical Zoology; Yakushiji 3311-1,
            Minamikawachimachi, Tochigi 329-0498, Japan
            (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339,
            Fax:81-285-44-6449)
FEATURES             Location/Qualifiers
     source          1..3189
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:2096"
     gene            1..3189
                     /gene="M11"
     CDS             1..3189
                     /gene="M11"
                     /codon_start=1
                     /transl_table=4
                     /product="120kDa membrane protein"
                     /protein_id="BAA74942.1"
                     /db_xref="GI:4240364"
                     /translation="MNISKKLKSYTLIGGLAVFGALGSASFGFKQSDKSNDNTQLVNQ
                     ARTLDANSVRLAGLGQNGSLFNTVLRDVDDNFITAANGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNRFDDRQTRAYYALLVNDEANVHLKRINTNSNRIGNRNNS
                     KFVIGGVDNPAHVIRFTDDGTKFNFTKQTGEIVNDFILDAPILPKDLHPDWYNLYIQ
                     RKILPDVNTAVVPWPVGRVSGNADDGMPFDFGNGQINTDPIAQTKTTDNPSTF
                     NSGAMPGANNRYDSQLNVKHRIKTSFQLDEKFVYPEWTGSEENKNITRLATGSLPSNE
                     RYWILDIPGTPQVTLKEDSVNVFSRLYLNSVNSLSFIGDSIYLFGTSELPSLWYYSFP
                     TRLSDLTALNQVKTDDIEASSTDNGTTNGTTTTADTSSGSTGAGTGNTTNTSQTVSN
                     PTLNTYRSFGIDSKPTSANKIDETNWADPNVIEARIYABYRLGIQNEIPITNAGNFIR
                     NTIGGVGFTSTGSRVVLRASYNGDQRPTGNEQPFLYVFGLGVQQTRTGTFWYGTYKL
                     LNNSPYDVDLDAARVGTETNQFRRTSLTYPVMGGYLTEBGARSFSNTPYIRAQGDTPES
                     RSIFQSGYSDNTYEYIQSVLGFDGIRNNLNVGVKASSPLNSNRPNPNGLEMIAATTYL
                     KSSVYTLRLADSSNPDASSSFSPTSLIDVNEIGVILPLLDNSFYTVNAAGNVALFSSN
                     PGSPGSYTAVNTFNQNLSDIAFBGSGAKYTSDFWGTIQFKPDEYLIQNGFTSQVARNF
                     VTNQSPLNSLVDFTPANAGTNRYRVVVDPDGNLTNQNLPLKVQIQYLDGKYYDAKLKNN
                     NLVTFSYNNFAALPSWVVPTAIGSTLGILAIMILGLAIGIPLRAQRKLQDKGFKTTF
                     KKVDTLTAAVGSVYKKIITQTANVKKKPAALGAGKSGDKKPAAAAKPAAPKSAPKA
                     SSPAKPTAPKSGAPTKPTAPKPAAPKPTAPKE"
ORIGIN
```

```
Query Match      98.6%;   Score 3145.8;   DB 1;   Length 3189;
Best Local Similarity   99.2%;   Pred. No. 0;
Matches 3162;   Conservative   0;   Mismatches   27;   Indels   0;   Gaps   0;
```

```
Qy       1 ATGAATATTTCTAAAAAAACTTAAAAGTTATAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA      60
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db       1 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA      60

Qy      61 GCTCTTGGTTCTGCAAGCTTGGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGTAACACCGCAA     120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCTCTTGGTTCTGCAAGCTTGGCTTTGGCTTTAAGCAATCAGATAACGTAACACGCAA     120

Qy     121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA     180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA     180

Qy     181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT     240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT     240

Qy     241 AATGGAACAATTATCAAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGT     300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AATGGAACAATTATCAAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGT     300

Qy     301 GATGATTGGTGATGACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGACAGAAAT     360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GATGATTTGCTGGATGACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGACAGAAAT     360

Qy     361 AGATTTGATCAAGACAAGACAAGAGCATATATTATGCTCGTGGTTAATGATGAAGCTAAC     420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGATTTGATCAAGACAAAGACAAGAGCATATATTATGCTCGTGGTTAATGATGAAGCTAAC     420

Qy     421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGTTAATAGAACAACAATTCT     480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGTTAATAGAACAACAATTCT     480

Qy     481 AAGTTTGTAATTGGTGGTGTGTTGATAATCCAGCTCACGTAATTACTTACTGATGATGGGG     540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AAGTTTGTAATTGGTGGTGTGTTGATAATCCAGCTCACGTAATTACTTACTGATGATGGGG     540

Qy     541 ACTAAATTTAATTTACAAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT     600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 ACTAAATTTAATTTACAAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT     600

Qy     601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTATATCATTTCAAAGAAAG     660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTATATCATTTCAAAGAAAG     660

Qy     661 ATCTTACCCAAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA     720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ATCTTACCCAAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA     720

Qy     721 ACAAATGCTGATGATGATGTTGATTGTGGAATGGTCAAATAACTAATACAGATCCT     780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ACAAATGCTGATGATGATGTTGATTGTGGAATGGTCAAATAACTAATACAGATCCT     780

Qy     781 AATTGCTCGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA     840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 AATTGCTCGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA     840

Qy     841 ATGCCTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA     900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 ATGCCTGGTCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA     900

Qy     901 TCTTTCCAATTGAGTAGATGAAAAATTTGTTTTATCCGAATCAACGATTGGTTCTGAAGAGAATAAA     960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TCTTTCCAATTGAGTAGATGAAAAATTTGTTTTATCCGAATCAACGATTGGTTCTGAAGAGAATAAA     960

Qy     961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC     1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC     1020
```

```
Qy    1021 ATACCCGGGACTCCACAAGTACTTTAAAAGAAGATTCAGTTAAGTATTTCAAGACTA     1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ATACCCAGGGACTCCACAAGTACTTTAAAAGAAGATTCAGTTAAGTATTTCAAGACTA     1080

Qy    1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTATTTATATATTTTGGTACC     1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATAGTATTTATATATTTTGGTACT     1140

Qy    1141 TCTGAATTACCATCATTATGTACTATTCATTCCAACTAGATTATCTGATCTAACCGCT     1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TCTGAATTACCATCATTATGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT     1200

Qy    1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA     1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA     1260

Qy    1261 AACGGAACAACCACACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT     1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 AACGGAACAACGACACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACACGGAAAT     1320

Qy    1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTAAATACTTAAATACTCGTAGTTTTGGA     1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ACTACTAACACTTCTCAAACAGTTTCTAATCCTAAATACTTAAATACTCGTAGTTTTGGA     1380

Qy    1381 ATTGATGATAAACCAACTTCTGCAAACAAAATAGATGAAACTAATGGGACAGATCCTAAC     1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 ATTGATGATAAACCAACTTCTGCAAACAAAATAGATGAAACTAATGGGACAGATCCTAAT     1440

Qy    1441 GTTATTGAAGCAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA     1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GTTATTGAAGCAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA     1500

Qy    1501 ACTAATGCAGGAAACTTATCCGAAACACAATTGTGGTGGTTGTTTTACTTCAACAGGT     1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ACTAATGCAGGAAACTTATCCGAAACACAATTGTGGTGGTTGTTTTACTTCAACAGGT     1560

Qy    1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAGAGCTTCCAACTGACTCAACGTCCAACTTCCAA     1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAGAGCTTCCAACTGACTCAACGTCCAACTTCCAA     1620

Qy    1621 CCTTTCTTATACGTATTTGGTTATTGGAACCAACAACTAGAACACAGAAACTTTCTGG     1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 CCTTTCTTATACGTATTTGGTTATTGGAACCAACAACTAGAACACAGGAACTTTCTGA     1680

Qy    1681 TACGGAACATATAAGCTTTTAAACACACAGCCCTTACGACGTATTAGAATTCTCCAAGAGTA     1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TACGGAACATATAAACTATTAAACACTATAAACACAGCCCTTACGACGTATTAGAATTCTCCAAGAGTA     1740

Qy    1741 GGTACTGAAACCAATCAATTTAGAGAGAACTTCATTCAACATCAATAACTAACATCCCTGTTA     1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GGTACTGAAACCAATCAATTTAGAGAGAACTTCATTCAACATCAATAACTAACATCCCTGTTA     1800

Qy    1801 CTAACTGAAAGAAGGTGCTAGAAGTTTCTCTAATACTTCCTAATAACTAATACAGATCCT     1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 CTAACTGAAAGAAGGTGCTAGAAGTTTCTCTAATACTTCCTAATAACTAATACAGATCCT     1860

Qy    1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT     1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT     1920

Qy    1921 CAATCAGTTTTAGGATTTGATGGAATTAGGAACTAACTTAAAGTGTGGGTTAAAGCATCA     1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CAATCAGTTTTAGGATTTGATGGAATTAGGAACTAACTTAAAGTGTGGGTTAAAGCATCA     1980

Qy    1981 AGCTTCTTAAACTCAAATGAACTAGAAATCAAAACGGTCTAGAAAATGATTGCTCAACAACA     2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 AGCTTCTTAAACTCAAATGAACTAGAAATCAAAACGGTCTAGAAAATGATTGCTCAACAACA     2040

Qy    2041 TACTTAAGATCACCAAGTTGGATTAGCTAGACAACATCTGGAATTACCCAAACCAACACCATTC     2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 TACTTAAGATCACAAATTGGATTAGCTAGACAACATCTGGAATTACCCAAACCAACCATTC     2100

Qy    2101 GGAACAACTCACCCAAGTTATTTCAGTATCACCTGTGATCATCAGTTCTCATCAATTAAGAAT     2160
```

Db   3181 AAAGAATAA 3189

RESULT 3
AE016967
LOCUS   AE016967   301042 bp   DNA   linear   BCT 03-SEP-2003
DEFINITION Mycoplasma gallisepticum strain R section 1 of 4 of the complete
            genome.
ACCESSION   AE016967 AE015450
VERSION   AE016967.1 GI:31541048
KEYWORDS   .
SOURCE   Mycoplasma gallisepticum R
  ORGANISM Mycoplasma gallisepticum R
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 301042)
  AUTHORS   Papazisi,L., Gorton,T.S., Kutish,G., Markham,P.F., Browning,G.F.,
            Nguyen,D.K., Swartzell,S., Madan,A., Mahairas,G. and Geary,S.J.
  TITLE   The complete genome sequence of the avian pathogen Mycoplasma
            gallisepticum strain R(low)
  JOURNAL   Microbiology (Reading, Engl.) 149 (Pt 9), 2307-2316 (2003)
  PUBMED   12949158
REFERENCE   2 (bases 1 to 301042)
  AUTHORS   Geary,S.J., Papazisi,L., Kutish,G., Mahairas,G., Swartzel,S.,
            Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,
            Mustafa,K. and Liao,X.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-OCT-2002) Department of Pathobiology and Veterinary
            Sciences, and Center of Excellence for Vaccine Research, The
            University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT
            06269-3089, USA
FEATURES   Location/Qualifiers
     source   1..301042
                     /organism="Mycoplasma gallisepticum R"
                     /mol_type="genomic DNA"
                     /strain="R"
                     /db_xref="taxon:233150"
     rep_origin   1..1889
                     /note="putative oriC"
     protein_bind   complement(24..32)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     protein_bind   complement(36..44)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     pulmonis consensus ttatcMaMa"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     protein_bind   114..122
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     repeat_region   209..217
                     /note="oriC imperfect repeat"
     protein_bind   complement(254..262)
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     mycoides consensus YtWtMHaMR"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     protein_bind   298..306
                     /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                     binding site; consensus sequence approach; Mycoplasma
                     mycoides consensus YtWtMHaMR"
                     /bound_moiety="DnaA"
                     /function="initiation of the chromosome replication"
     repeat_region   complement(401..450)
                     /note="oriC imperfect repeat"
     repeat_region   complement(462..510)
                     /note="oriC imperfect repeat"

Qy   2101 GGAACAACTCACCACCAAGTTATTTCAGTAGTAAACCAGTTATGGTACTTCTTATTCAGTATCATCAATTCATCAATTCAT 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCTTATTCACAAATGAAAATAAT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCTTATTCACAAATGAAAATAAT 2220

Qy   2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGGCGTCAAGCTCA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AAATCTAGTGTTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280

Qy   2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGAATCTTACCTTTATTAGAC 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGAATCTTACCTTTATTAGAC 2340

Qy   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCAAACCCTGGT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCAAACCCTGGT 2400

Qy   2401 TCTCCTGGTGATCATATACTGCTGTAAATACATTAATCAGAACTATCTGATATTGCTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460

Qy   2461 GAAGGTTCTGGTGCTCAAGTATACATCTGATTCTGGGGAACAATCCAATTCAAACCCGAT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GAAGGTTCTGGTGCTCAAATATACATCTGATTCTGGGGAACAATCCAATTCAAACCCGAT 2520

Qy   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580

Qy   2581 AGCTTCTTAAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 AGCTTCTTAAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy   2641 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GTGGTTGATCCTGATGGGTAATTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700

Qy   2701 TACTTAGATGGTAAGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TACTTAGATGGTAAGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760

Qy   2761 TATAACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TATAACAACTTTGCCTTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Qy   2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA 2880

Qy   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT 2940

Qy   2941 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000

Qy   3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCGCTGCTGCTAAACCTGCT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCGCTGCTGCTAAACCTGCT 3060

Qy   3061 GCTCCAGCTAAACCATCTGCACCAAAAAGCTAGCTCACCCAGCTAAACCAACTGGGCCTAAA 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCTCCAGCTAAACCATCTGCACCAAAAAGCTAGCTCACCCAGCTAAACCAACTGCGCCTAAA 3120

Qy   3121 TCTGGTGCGCGCTACAAACCAACTGCTCCTCCAAAACCAACCGCTCCC 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TCTGGTGCGCGCTACAAACCAACTGCTCCTCCAAAAACCAACCGCTCCC 3180

Qy   3181 AAAGAATAA 3189
          |||||||||

```
repeat_region   complement(575..623)
                /note="oriC imperfect repeat"
repeat_region   complement(713..761)
                /note="oriC imperfect repeat"
repeat_region   complement(763..811)
                /note="oriC imperfect repeat"
repeat_region   complement(821..869)
                /note="oriC imperfect repeat"
repeat_region   872..920
                /note="oriC imperfect repeat"
repeat_region   991..1004
                /note="oriC imperfect repeat"
repeat_region   1061..1109
                /note="oriC imperfect repeat"
repeat_region   1147..1195
                /note="oriC imperfect repeat"
protein_bind    complement(1185..1192)
                /note="putative DnaA-box; Mycoplasma spp. putative dnaA
                binding site; consensus sequence approach; Mycoplasma
                capricolum consensus ttWtMÏaa"
                /bound_moiety="DnaA"
                /function="initiation of the chromosome replication"
repeat_region   complement(1230..1278)
                /note="oriC imperfect repeat"
repeat_region   1326..1374
                /note="oriC imperfect repeat"
repeat_region   complement(1551..1564)
                /note="oriC imperfect repeat"
repeat_region   complement(1609..1657)
                /note="oriC imperfect repeat"
repeat_region   complement(1700..1748)
                /note="oriC imperfect repeat"
repeat_region   complement(1831..1879)
                /note="oriC imperfect repeat"
gene            1890..2681
                /gene="parA"
                /locus_tag="MGA_0619"
                /synonym: soj"
CDS             1890..2681
                /gene="parA"
                /locus_tag="MGA_0619"
                /note="MGA_0619 ATPase involved in chromosome
                partitioning; pfam00991 ParA family ATPase; COG1192 Soj
                ATPases involved in chromosome partitioning, MGR_001"
                /codon_start=1
                /transl_table=4
                /product="ParA/Soj"
                /protein_id="AAP56351.1"
                /db_xref="GI:31541049"
                /translation="MMIISFINNKGGVLKTTLATNICGVFSKFFPKSRSVIVDLDGQG
                NVSATFGQHPERLKNTLDIFRGEKDIDDCVLNVFPSIDILPSNHELSFVDMDVARKE
                YKLSVIKNLIEKLEEMYDPVFLDTPPAMSTIVSVSMHISDMIVIPFEPDQYSILGLLR
                VIDTIETFPRENEDLKVLVVPTKVNARTLHNDVLNIVKTKLSKKNIPLSKNIVSFTT
                KSSASVGYEKLPIVLINQRSKYQEYIEITKEIINLLKDNKNNEK"
gene            2659..3147
                /locus_tag="MGA_0621"
CDS             2659..3147
                /locus_tag="MGA_0621"
                /note="MGA_0621 PS50318; MGR_002"
                /codon_start=1
                /transl_table=4
                /product="unique hypothetical"
                /protein_id="AAP56352.1"
                /db_xref="GI:31541050"
                /translation="MTRTMKNKAKKKERRFTDLSADLDEEVEKIDPEYEDRDKEIKIE
                KNKDNQVIDKNDPFYSESFEEARIQLIKDKKYEVKKEEBKVQETTVKNKISEAKKEE
                AKDVYIDSSLEIASQEPLTKGMHFYTNSRIIRKVRBCAKNKGLSISRLITMILDKSIK
                EE"
gene            3163..4548
                /gene="dnaA"
                /locus_tag="MGA_0622"
CDS             3163..4548
                /gene="dnaA"
```

```
                /locus_tag="MGA_0622"
                /note="MGA_0622 ATPase involved in DNA replication
                initiation COG0593"
                /codon_start=1
                /transl_table=4
                /product="DnaA"
                /protein_id="AAP56353.1"
                /db_xref="GI:31541051"
                /translation="MKTKLKRFLEISVHFNEANSELLDAFVHSIDFVFBENDNIYIY
                PESPYYFPNEFKNKLNHLINVENAVVFNDYLSLEWKCIIKSNKRVNLLNKKEADTLKEK
                LATLKKQEKYKINPLSKGIKEKYNFGNYLVFEFNKEAVYLAKQIANKTTHSNWNPIII
                EGKPGVGKSHLLQAIANERQKLFPEBKICVLSSDDFGSBFIKSVIAPDTHIESFGSK
                YKDYDLLMIDDVQIISNRPKTNETPFTIFNSLVDQKKTIVITLDCKIBEIQDKLTARM
                ISRFQKGINVRINQPKNEIIQIFKQKFKENNLEKYMDDHVIEEISDFDEGDIRKIEG
                SVSTLVFMNQMYGSTKTKDQLKSFIEKVTNRKNJILSKDPKYVFDKIKYHFNVSEDV
                LKSSKRKKEIVQARHICMVYLKNVVNKNLSQIGKLLRKDHTTVRHGIDKVEBELENDP
                NLKSFLDLFKN"
gene            4744..5133
                /locus_tag="MGA_1322d"
                /note="MGA_1322d"
CDS             4744..5133
                /locus_tag="MGA_1322d"
                /note="Q COG1132 ABC-type multidrug/protein/lipid
                transport system, ATPase component; similar to MGA_1284;
                MGR_004"
                /codon_start=1
                /transl_table=4
                /product="ABC transporter component"
                /protein_id="AAP56354.1"
                /db_xref="GI:31541052"
                /translation="MNRKLDHCISALLYFGLINFGIIGRDKNALTQYSLILLGVAAVG
                PTFGYLGGRSIIASVEFAKQLRVNIFERYQSFSVKNTDKFEKASVLTRMTTDINFIH
                QSIQSGRTAIRGMSVPLFSLVLMFVTS"
gene            5130..6077
                /locus_tag="MGA_0625"
                /note="synonym: mldB"
CDS             5130..6077
                /locus_tag="MGA_0625"
                /note="MGA_0625, MdlB-like; ABC-type
                multidrug/protein/lipid transport system MdlB [Q] COG1132
                PS00890; similar to MGA_1285, MGA_0626 and MGA_1287;
                MGR_005"
                /codon_start=1
                /transl_table=4
                /product="MdlB-like"
                /protein_id="AAP56355.1"
                /db_xref="GI:31541053"
                /translation="MKLGIASVAIMPVIIGGILLIVRFVIGNYKKLPKQYDQLNNLAK
                ESIAGARVVKSTHQQDNEIQKPNRVAGFIYKNFTKIERITALISPIVLFCIYALAIAI
                AWIGTNNIVDGKLDIGSLASVFAYAFQMLINLLLSVVYVTIITAKPSKDRIIEVLTE
                KIDIKDKGVIKDIDSLASVFAYAFQMLINLLLSVVYVTIITAKPSKDRIIEVLTE
                KIDIKDKGVIDYDYEVKKSVPDNPHNLEKINIKIKKGQTIGIIGSTGS
                GKTSIVNLLTRLYECNEBGQVLLNNIQLNNYSIKALRDAIAIVPQKSILYSGTIKDNIL
                MGGNYSDEEVEKAITCAQAEFINKLPNKLDSIVEQKWN"
gene            6369..8294
                /locus_tag="MGA_0626"
CDS             6369..8294
                /locus_tag="MGA_0626"
                /note="MGA_0626, ABC-type multidrug/protein/lipid
                transport system; similar to MdlB [Q] COG1132 PS00890;
```

```
Query Match       98.6%;   Score 3144.2;   DB 1;   Length 301042;
Best Local Similarity 99.1%;   Pred. No. 0;
Matches 3161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy      1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATTAGGTGGATTAGCTGTATTTGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 226607 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATTAGGTGGATTAGCTGTATTTGA 226666

Qy     61 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATCAGATAAGAGTAACCATAACACGCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 226667 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATCAGATAAGAGTAACGACAATAACACGCCAA 226726

Qy    121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGGACAA 180
```

DNA sequence alignment (Db / Qy), rotated 90°.

Left column blocks (position indices):

Db	Qy
226727	226786
181	240
226787	226846
241	300
226847	226906
301	360
226907	226966
361	420
226967	227026
421	480
227027	227086
481	540
227087	227146
541	600
227147	227206
601	660
227207	227266
661	720
227267	227326
721	780
227327	227386
781	840
227387	227446
841	900
227447	227506
901	960
227507	227566
961	1020
227567	227626
1021	1080
227627	227686
1081	1140
227687	227746
1141	1200
227747	227806
1201	1260

Right column blocks (position indices):

Db	Qy
227807	227866
1261	1320
227867	227926
1321	1380
227927	227986
1381	1440
227987	228046
1441	1500
228047	228106
1501	1560
228107	228166
1561	1620
228167	228226
1621	1680
228227	228286
1681	1740
228287	228346
1741	1800
228347	228406
1801	1860
228407	228466
1861	1920
228467	228526
1921	1980
228527	228586
1981	2040
228587	228646
2041	2100
228647	228706
2101	2160
228707	228766
2161	2220
228767	228826
2221	2280
228827	228886
2281	2340
228887	228946

```
REFERENCE          Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
                   1 (bases 1 to 8354)
AUTHORS            Goh,M.S., Gorton,T.S., Forsyth,M.H., Troy,K.B. and Geary,S.J.
TITLE              Molecular and biochemical analysis of a 105 kDa Mycoplasma
                   gallisepticum cytadhesin (GapA)
JOURNAL            Microbiology 144 (Pt 11), 2971-2978 (1998)
MEDLINE            99061189
PUBMED             9846732
REFERENCE          2 (bases 1 to 8354)
AUTHORS            Papazisi,L., Troy,K.E., Gorton,T.S., Liao,X. and Geary,S.J.
TITLE              Analysis of cytadherence-deficient, GapA-negative Mycoplasma
                   gallisepticum strain R
JOURNAL            Infect. Immun. 68 (12), 6643-6649 (2000)
MEDLINE            20536401
PUBMED             11083776
REFERENCE          3 (bases 1 to 8354)
AUTHORS            Papazisi,L., Troy,K.B. and Geary,S.J.
TITLE              Direct Submission
JOURNAL            Submitted (09-DEC-1999) Department of Pathobiology, University of
                   Connecticut, 61 North Eagleville Rd. U-89, Storrs, CT 06268-3089,
                   USA
FEATURES           Location/Qualifiers
     source        1..8354
                   /organism="Mycoplasma gallisepticum"
                   /mol_type="genomic DNA"
                   /strain="R"
                   /db_xref="taxon:2096"
     gene          667..3756
                   /gene="gapA"
     CDS           667..3756
                   /gene="gapA"
                   /note="GapA; MGP1; cytadhesin; similar to Mycoplasma
                   pneumoniae P1"
                   /codon_start=1
                   /transl_table=4
                   /product="adherence protein A"
                   /protein_id="AAF25381.1"
                   /db_xref="GI:6694735"
                   /translation="MIWTFDPGNTNDLTGKVGFYDANNRLTAPSGDVPFNVSDLSSKT
                   VVEATDQEDPNVFYLLLIPDAAVQEQKTKDQVPENYFMSDAPATGDTSAEGSATPA
                   GGGSSSAACGGAVAPAAASSTARLVBEGNSAGMGTMTPTASTSETVIDYNSDQNKIP
                   KPKTLLDSSESSESSINGGRTYANINTQNNLQGVIVKVNENLFNSENPFAVENMAFIKP
                   KDMVDNYPSTWTQGSANGKMTNVLQCYKHDNPAVNNRFYRAKYFPKRLETQTTTPLI
                   DSSFSPYEHPEWYEDNQFVMPWMQYITNLGGLYAXDGMVLFGGNGTWVNNESALSIG
                   VFRTKFENRIAEAPGNTKTVGPYGYGILLSAISFDATRNGLALAPASLGQDVVHVPR
                   LAVGGVSSPRGANGNIPLFGSAITWGNGGNFLDTKWHSPAVIEDAPTTFVTVNSSGAL
                   QNSGNPQPTSTPMPNSNGNESIPYRWTNSYDYNSVRFAALISKPAGGNTKQVESLFTT
                   ALKLDTLNSLPNKFTQENNIFFSYAMLDGRQWSLGTRKOSAWLATNTINNPTYNTQQQ
                   LASTVAGENANPRNIINALTTAKGFDRRDIGNVVTYSNVTNNKFTYYQVGGAITTWP
                   EVQVNYKTSANITYNLTRTDPGSTTPATQDANTVSSKLNGAYLSSTGDQGWYNGSI
                   YVKKASFTPSSGYTWQDFKGLTTTASNAVISNWTKAGYSIRPDDDTVFNASKIPPEK
                   EITAAVNVRSLDSYYVQLNGETSVNTVARVSPDSSALALNPNRITNPLMNRDNVIGQG
                   AFISRNDIPSSFFENKINDIVTTEADGKEVLDSKYINSIYRYTPPQNNPDIRLRLLVI
                   DRSRATNDFIKLLPQVLVDGEYVAVPQANSVFVSDQEBFTGFDALPGVVLPVAISIPII
                   IIALALALGLGIGIPMSQNRKMLKQGFAISNKKVDILTTAVGSVFKQIINRTSVTNIK
                   KTPQMLQANKKDGASSPSKPSAPAAKKPAGPTKSPAPGAKPTAPAKPKPAPTKKIE"
     gene          3778..6966
                   /gene="crmA"
     CDS           3778..6966
                   /gene="crmA"
                   /note="CrmA; similar to Mycoplasma pneumoniae MGPC,
                   Mycoplasma genitalium MGPC, Mycoplasma gallisepticum
                   strain S6 GapA, Mycoplasma pirum P1-like adhesin,
                   Mycoplasma pneumoniae ADP1, and Mycoplasma genitalium
                   ADP1, encoded by GenBank Accession numbers Q50341, P22747,
                   AF063376, Q50371, AAB95661 and P20796, respectively"
                   /codon_start=1
                   /transl_table=4
                   /product="putative cytadherence related molecule A"
                   /protein_id="AAF25382.1"
                   /db_xref="GI:6694736"
                   /translation="MNISKKLKSYTLIGGLAVFGALGSASPGFKQSDKSNDNTQLVNQ
                   ARTLDANSVRIAGLGQNGSLFNTVLRDVDDNFITAANGTIIKLDSFTKPLYGLDLSDD
```

```
Qy    2341  AATTCATTCTATACAGTAAATGCTGCTGTAAATGTTGCATTGTTCTCATCATCAAACCCTGCT  2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  228947  AATTCATTCTATACAGTAAATGCTGCTGTAAATGTTGCATTGTTCTCATCATCAAACCCTGGT  229006

Qy    2401  TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT  2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229007  TCTCCTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT  229066

Qy    2461  GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGAACAATCCAATTCAAACCCGAT  2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229067  GAAGGTTCTGGTGCTAAATATACATCTGATTTCTGAGGAACAATCCAATTCAAACCCGAT  229126

Qy    2521  GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA  2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229127  GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA  229186

Qy    2581  AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA  2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229187  AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA  229246

Qy    2641  GTGGTTGATCCTGATGGTAATTAACAAACAAAACCTACCTCTAAAGTTCAGATCCAA  2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229247  GTGGTTGATCCTGATGGTAATTAACAAACAAAACCTACCTCTAAAGTTCAGATCCAA  229306

Qy    2701  TACTTAGATGTGTAAGTATTATGATGCTAAATTAAAGAACAATATATTTAGTAACATTCTCT  2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229307  TACTTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATATATTTAGTAACATTCTCT  229366

Qy    2761  TATAACAACTTTGGCGCCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA  2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229367  TATAACAACTTTGGCGCCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA  229426

Qy    2821  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA  2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229427  GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAA  229486

Qy    2881  AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT  2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229487  AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT  229546

Qy    2941  GCTCGTTGGTTCAGTTGTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT  3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229547  GCTCGTTGGTTCAGTTGTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT  229606

Qy    3001  GCTGCTTAGTGCTGGTAAATCTGTGATAAGAAACCTGCTCGTCTGCTAAACCTGCT  3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229607  GCTGCTTAGTGCTGGTAAATCTGTGATAAGAAACCTGCTCGTCTGCTAAACCTGCT  229666

Qy    3061  GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGCCTAAA  3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229667  GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGCGCCTAAA  229726

Qy    3121  TCTGGTGCGGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC  3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  229727  TCTGGTGCGGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC  229786

Qy    3181  AAAGAATAA  3189
            |||||||||
Db  229787  AAAGAATAA  229795
```

```
RESULT 4
AF214004
LOCUS       AF214004      8354 bp   DNA      linear    BCT 22-NOV-2000
DEFINITION  Mycoplasma gallisepticum adherence protein A (gapA), putative
            cytadherence related molecule A (crmA), and putative cytadherence
            related protein B (crmB) genes, complete cds.
ACCESSION   AF214004
VERSION     AF214004.1  GI:6694734
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
```

```
        PAGYKVKQIVSDYTTSRNRFDQRQTRAYYALLVNDEANVHLKRINTNSNRIGNRNNS
        KFVIGGVDNPAHVIRFTDDGTKFNFTKQTGEIVNDFILDAPILPKDLLHPDWTNLYIQ
        RKILPNDVNTAVVPWRGVRSGTNADDGMFDFGNGQITNTDPIAQTKTTDNQNPSTF
        NSGAMPGANNRYDSQLNVKHRIKTSFQLDEKFVYPEWTGSEENKNITRLATGSLPSNE
        RYWILDIPGTPQVTLKEDSVNVFSRIYLNSVNSLSFIGDSIYIFGTSBLPSLWYYSFP
        TRLSDLTALNGVKTDDIEASSTDNGTTNGTTTDTSSSGTSAGTGNTTNTSQTVSN
        PTLNVTYRSFGIDSKPTSANKIDETNWADPNVIEARIYAEYRLGIQNEIPITNAGNFIR
        NTIGGVGFTSTGSRVVLRASYNGDQRPTGNFQPFLYVFGVIGLGYQQTRIGTFWYGYKL
        LNNSPYDVLDDARVGTETNQFRRTSLTYPYMGGVLTEBGARSFSNTPYIRAQGDTPES
        RSIFOSGYSDNTYEYIQSVLGFDGIRNNLNVGVKASSFLNSNRPNPNGLEMIAATTYL
        RSQIGLARTSGLPNQQPFGTTHQVISVSPGDQFSSIKNIRTIFPGNQLWFPLFTNENN
        KSSVVTLRLADSSNPDASSSFPTSLLDVNEIGVLPLLDNSFYVNAAGNVALFSSN
        PGSPGSYTAVNTFNQNLSDIAFEGSGAKVTSDFWGTIQFKPDEVLIQNGFSTSQVARNF
        VTNQSFLNSLVDFTPANAGTNYRVVVDPDGNLTNQNLPLKVQIQYLDGKYYDAKLKNN
        NLVTFSYNNFRALPSWVVPTAIGSTLGILAIMIILGLAIGIPLRAQRKLQDKGFKTTF
        KKVDTLTAAVGSVYKKIITQTANVKKPAALGAGHKSGDKKPAAAKPAAPAKPSAPKA
        SSPAKPTAPKSGAPTKPTAPKPAAPKTAPKE"
```

gene	7128..8126	
	/gene="crmB"	
CDS	7128..8126	
	/gene="crmB"	
	/note="CrmB; similar to N-terminus of Mycoplasma gallisepticum GapA"	
	/codon_start=1	
	/transl_table=4	
	/product="putative cytadherence related protein B"	
	/protein_id="AAF25383.1"	
	/db_xref="GI:6694737"	
	/translation="MLKAHIKRSLIFILLIAAISLIFIGLFGFKKPNINQVKLNNQ	
	SSSLITNQPSDQIIRKYSLDSIRWNANANFSSSPTKDGGLLVITTDGRLERIDSPGYL	
	KWELNLRNINALATVDEEQNLVYSDNRIVIKPLDAVFINNVSKKVVAERAIQSQEDPTI	
	YYLLAVSGGLITNDLKDQLLSFQQLATDLTSQGVVLKVKENPTQISDPVGIIDYAMIS	
	PTDLTQNYPASWKTNTLSLYGDNGSFRKADHPSWFNNDSNFISLPWLQYITNLANLYB	
	YKDNVFLFGGNGNWVNAIRVFVSKQPNTFLLGCLESGFIKILVLDLSQVILMLTY"	

ORIGIN

```
Query Match      98.6%;  Score 3143.4;  DB 1;  Length 8354;
Best Local Similarity 99.1%;  Pred. No. 0;  Indels  28;  Gaps  0;
Matches 3159;  Conservative  2;  Mismatches  2;
```

```
Db  5338 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA 5397

Qy  1621 CCTTTCTTATACGTATTGGTTATTTAGGATACCAACAAACTAGAACAGGAACTTCTCTGG 1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5398 CCTTTCTTATACGTATTTGGTTATTTAGGATATCCAACAAACTAGAACAGGAACTTCTCTGA 5457

Qy  1681 TACGGAACATATAAGCTTTAAACAACACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5458 TACGGAACATATAAACATTAAAACATTAAAAACTAGAACGTATTAGATGCTGCAAGAGTA 5517

Qy  1741 GGTACTCTGAAACCAATCAATTCAATTCATTAACACCTGTATGCGGGTGGATAT 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5518 GGTACTCTGAAACCAATCAATTCAATTTAGAAGAACTTCATTAACATCCTGTATGCGGGTGGATAY 5577

Qy  1801 CTTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAGAGACCAAGGTGAC 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5578 CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGACCAAGGTGAC 5637

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGCTATTCTGTAATAACTTATGAGTACATT 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5638 ACACCAGAAAGCCGAAGCATCTTCCAATCTGCTAYTCTGTAATAACTTATGAGTACATT 5697

Qy  1921 CAATCAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5698 CAATCAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA 5757

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5758 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA 5817

Qy  2041 TACTTAAGATCACAAAATTGGATTAGCTAGACATCTGGATTACCAAACCAACAACCATTC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5818 TACTTAAGATCACAAAATTGGATTAGCTAGACATCTGGATTACCAAACCAACCAACCATTC 5877

Qy  2101 GGAACAACTCACCAAGTTATTCAGTATCACCTGGTGATCAGTTCTCATCAATTAAGAAT 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5878 GGAACAACTCACCAAGTTATTTCAGTATCACCTGGTGATCAGTTCTCATCAATTAAGAAT 5937

Qy  2161 ATTAGAAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAAATGAAAATAAT 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5938 ATTAGAAACAATCTTCCCTGGTAACCAGTTATGATACTTCTTATTCACAAAATGAAAATAAT 5997

Qy  2221 AAATCTAGTGTTTATACATTAAGATTTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAGCTCA 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5998 AAATCTAGTGTTTATACATTAAGATTTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAGCTCA 6057

Qy  2281 TTCAGTCCAACAGTTTAATTGACGTTAATGACGTTAATCTTACCTTATTAGAC 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6058 TTCAGTCCAACAGTTTAATTGACGTTAATGACGTTAATCTTACCTTTATTAGAC 6117

Qy  2341 AATTCATTCATAGTTCTATACAGTAGTGCTGCTGCGTGTAAATGCTGGTGGTCATTGTTGTCATGTTGCAATGCATGCATTCATCAAACCCTGGT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6118 AATTCATTCATAGTTCTATACAGTAGTGCTGCTGCGTGTAAATGCTGCTGGTTCATTGTTGTTCATGTTGCAATTCATCAAACCCTGGT 6177

Qy  2401 TCTCCTGGATCATATACTGCTGTAAATACTGCTGTAAATACTTTAATCAGAACTTATCTGATATTGCTTTT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6178 TCTCCTGGATCATATACTGCTGTAAATACTGCTGTAAATACTTTAATCAGAACTTATCTGATATTGCTTTT 6237

Qy  2461 GAAGGTTCTGGTGCTAAGTATCATACATCTGATTTCTGGGGAACAATCCAAATCCAAATTCAAACCCGAT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6238 GAAGGTTCTGGTGCTAAGTATCATACATCTGATTTCTGGGGAACAATCCAAATCCAAATTCAAACCCGAT 6297

Qy  2521 GAGTACTTAATTCAAAATGGGGTTCACCAGGTTCAGGTGGCTAGAAACTTCGTTACGAAACCTGGT 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6298 GAGTACTTAATTCAAAATGGGGTTCACCAGGTTCACGTGGCTAGAAACTTCGTTACGAAACCAA 6357

Qy  2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTTCACTCCTGCTAATGCTCTGGTACTAACTACCGTGTA 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6358 AGCTTCTTAAACAGTTTAGTTGACTTCACTTCACTCCTGCTAATGCTCTGGTACTAACTACCGTGTA 6417

Qy  2641 GTGGTTGATCCTGATGGTGCTAAGTTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6418 GTGGTTGATCCTGATGGTGCTAAGTTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 6477
```

```
Db  2701 TACTTAGATGGTAAGTAATTATATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6478 TACTTAGATGGTAAGTATTATATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTCT 6537

Qy  2761 TATAACAACTTGCGCGCCTTACCTCATGGGTATGCCTACAGCAAATGGTAGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6538 TATAACAACTTGCGCGCCTTACCTCATGAGTATGCCTACAGCAAATGGTAGTAGTACATTA 6597

Qy  2821 GGTATTCTTGCAATTATGATCATCTTAGGATATCGGTATTCCTTAAGAGCTCAA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6598 GGTATTCTTGCAATTATGATCATCTTAGGATATCGGTATTCCTTAAGAGCTCAA 6657

Qy  2881 AGAAAATACAAGACAAAGGGTCAAACAACATTCAAAAAAGTGATACCTGACTGCT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6658 AGAAAATACAAGACAAAGGGTTCAAACAACATTCAAAAAAGTGATACCTTGACTGCT 6717

Qy  2941 GCTGTTGGTTCAGTTACAAGAGATTATTACCCAAACTGCTACGTTAAGAAAAACCT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6718 GCTGTTGGTTCAGTTACAAGAGATTATTACCCAAACTGCTACGTTAAGAAAAAACCT 6777

Qy  3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGTAAGAAACCTGCTGCTGCTGCTAAACCTGCT 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6778 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGTAAGAAACCTGCTGCTGCTGCTAAACCTGCT 6837

Qy  3061 GCTCCAGCTAAAACCATCTGCACCCAAAAGTAGCTCACCAGCTAAACCCAACTGGGCCTAAA 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6838 GCTCCAGCTAAACCATCTGCACCCAAAAGCTAGTCACCAGCTAAACCCAACTGCGCCTAAA 6897

Qy  3121 TCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACCGCTCCC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  6898 TCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACCGCTCCC 6957

Qy  3181 AAAGAATAA 3189
         |||||||||
Db  6958 AAAGAATAA 6966
```

```
RESULT 5
AX113685

LOCUS        AX113685            10651 bp    DNA       linear   PAT 01-MAY-2001
DEFINITION   Sequence 1 from Patent WO0127254.
ACCESSION    AX113685
VERSION      AX113685.1   GI:13939855
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    Schwarzmann,F.
  TITLE      Gene transfer vectors for treating autoimmune diseases and diseases
             with immunopathogenesis by therapy
  JOURNAL    Patent: WO 0127254-A 1 19-APR-2001;
             Schwarzmann, Fritz (DE)
FEATURES             Location/Qualifiers
     source          1..10651
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Plasmid"
ORIGIN

Query Match       98.6%;  Score 3143.4;  DB 6;  Length 10651;
Best Local Similarity  99.1%;  Pred. No. 0;
Matches 3159;  Conservative  2;  Mismatches  28;  Indels  0;  Gaps  0;

Qy    1 ATGAATATTTCTAAAAAACTTAAAAACTTAAAAGTTATACATTGATAAGTGGATTAGGTGATTAGCTGTGATATTTGA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db    6 ATGAATATTTCTAAAAAACTTAAAAACTTAAAAGTTATACATTGATAAGTGGATTAGGTGATTAGCTGTGATATTTGA 65

Qy   61 GCTCTTGGTTCTGCAAGTTTGGCTTTAAGCAATCAGATAACGTAACGATAACACGCAA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 GCTCTTGGTTCTGCAAGCTTTGCAAGTTTCAGCAATCAGATAACGTAACGATAACACGCAA 125
```

```
Qy   121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGTTAGACTTGCAGGTCTTGGACAA  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   126 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGTTAGACTTGCAGGTCTTGGACAA  185

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   186 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  245

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   246 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  305

Qy   301 GATGATTGTGGTGGATACAAAGTAAAACAAAGTTTCGATACTACAACTAGACAGAAAT  360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   306 GATGATTTGCTGGATACAAAGTAAAACAAAGTTTCAGATTACAACTAGACAGAAAT  365

Qy   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGAATGAAGCTAAC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   366 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGAATGAAGCTAAC  425

Qy   421 GTTCATTTAAAAAGAATTAATAACTAACTCAAATAGGTAATAGAAAACAACAATTCT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 GTTCATTTAAAAAGAATTAATAACTAACTCAAATAGGTAATAGAAAACAACAATTCT  485

Qy   481 AAGTTGTAATTGGTGGTGTTGATATCCAGCTCACGTAATTAGAGTTACTGATGATGG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 AAGTTTGTAATTGGTGGTGTTGATATCCAGCTCACGTAATTAGAGTTACTGATGATGGG  545

Qy   541 ACTAAATTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   546 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  605

Qy   601 GCGCCAATCTTACCTAAAGATTGACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 GCGCCAATCTTACCTAAAGATTGACACCCAGATTGATATAACTTATACATTCAAAGAAAG  665

Qy   661 ATCTTACCACAAGACTCAACACTCGCAGTTGTTCCTTGGCCAGTAGAGTTAGTGGA  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 ATCTTACCACAAATGACGTAAACACTGCAGTTGTTCCTTGACCAGTAGAGTTAGTGGA  725

Qy   721 ACAAATGCTGATGATGGGAATGTTGATTGTGTGGGAATGGTCAAATAACTAATAACAGATCCT  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 ACAAATGCTGATGATGGGAATGTTGATTGTTTGATTTTGGGAATGGTCAAATAACTAATAACAGATCCT  785

Qy   781 ATTGCTCAAACTAACCACTACTGATAAATCAAAATCCTTCAACTTTTAATTCAGGAGGCA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 ATTGCTCAAACTAACCACTACTGATAAATAATCAAAATCCTTCAACTTTTAATTCAGGAGGCA  845

Qy   841 ATGCCTGGTGCAAACAAATGAATAGATACGATTCTCAATTGTCAAGCATAGAATTAAAACA  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 ATGCCTGGTGCAAACAAATGACGATTCTCAATTGTCAAGCATAGAATTAAAACA  905

Qy   901 TCTTTTCCAATTGAATGAAAAAATTTGTTTATCCGAAATGACTCGTTCTGAAAGAGAATAAA  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 TCTTTTCCAATTGAATGAAAAAATTTGTTTATCCGAAATGACTCGGTTCTGAAAGAGAATAAA  965

Qy   961 AATAATTACAAGACTCCACCAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 AATAATTACAAGACTCCACCAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA  1025

Qy  1021 ATACCCGGGACTCCACAAGTTAAAGAAGATTTAACGTATTTTATATTTTGGTACC  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026 ATACCAGGGACTCCACAAGTTAAAGAAGATTTAACGTATTATATTTTGGTACC  1085

Qy  1081 TACTTAAACTCAGTTAATTCTTATCATTCATTGTGATGTATTTATATTTTTGGTACT  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086 TACTTAAACTCAGTTAATTCTTATCATTCATTGGTGATTAGTATTTATATTTTTGGTACT  1145

Qy  1141 TCTGAATTACCATCATTATGGTACACTAGATTCATTCCCCAACTAGATCTGATCTAACCGCT  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1146 TCTGAATTACCATCATTATGATGATTATCTGATTCATTCCCCAACTAGATCTGATCTAACCGCT  1205


Qy  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1206 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1265

Qy  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1266 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1325

Qy  1321 ACTACTAACACTTCTCAAACAGTTCTATCCTAATCTTAAATACTTATCGTAGTTTTGGA  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1326 ACTACTAACACTTCTCAAACAGTTCTATCCTAATCTTAAATACTTATCGTAGTTTTGGA  1385

Qy  1381 ATTGATAGTAACCAACTTCTGCAAACAAAATAGATGAAATCAATTGGGCAGATCCTAAC  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1386 ATTGATAGTAAACCAACTTCTGCAAACAAAAATGATGAAACTAATTGAGCAGATCCTAAT  1445

Qy  1441 GTTATTGAAGCAGAATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1446 GTTATTGAAGCAGAATATGCTGAATACAGATTAGTATTCAAAATGAAATTCCAATA  1505

Qy  1501 ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTGTTTACTTCAACAGGT  1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1506 ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTGTTTACTTCAACAGGT  1565

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAAGTCGTCAACTGGTCAACAGGAACTTCCAA  1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1566 TCAAGAGTAGTTTTAAGAGCTTCTTATAAGTCGTCAACGTGTCAACTGGAAACTTCCAA  1625

Qy  1621 CCTTTCTTATACGTATTGGTTATTTAGGATGTGAAATTGTTAATGACCTTCATTTTAGAT  1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1626 CCTTTCTTATACGTATTGGTTATTTAGGATGTGAAATTGTTAATGACCTTCATTTTCTGA  1685

Qy  1681 TACGGAACATATAAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1686 TACGGAACATATAAACTATTAAACAACAGCCCTTACGACGTATTAGATGCTGCAAGAGTA  1745

Qy  1741 GGTACTGAAAACCAATCAATCAATTAGAAGAACTTCATTAACATACACCCTGTTATGGGTGGATAT  1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1746 GGTACTGAAAACCAATCAATCAATTAGAAGAACTTCATTAACATACACCCTGTTATGGGTGGATAY  1805

Qy  1801 CTAACTGAAGAAGTGCTAGAAGTTCTCTAATAACTAATAACTCCATATAAGAGCACAAGGTGAC  1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1806 CTAACTGAAGAAGTGCTAGAAGTTCTCTAATAACTAATAACTCCATATAAGAGCACAAGGTGAC  1865

Qy  1861 ACACCAGAAAGCCGAAGCATCTCCAATCTGGCTACTCGATAATACTTATGAGTACATT  1920
         ||||||||||||||||||||||||  :|||||||||||||||||||||||||||||||
Db  1866 ACACCAGAAAGCCGAAGCATCTCCAATCTGGCTAYTCGATAATACTTATGAGTACATT  1925

Qy  1921 CAATCAGTTTGATTTTGATGGAATTAGAAATAACTAAAATGTTGGGGTTAAAGCATCA  1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1926 CAATCAGTTTGATTTTGATGGAATTAGAAATAACTAAAATGTTGGGGTTAAAGCATCA  1985

Qy  1981 AGTTCTTAAACTCAATAGACCAAATCCAACGGTCTAGAAATGATTGCTGCAACAACA  2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1986 AGCTTCTTAAACTCAATAGACCAAATCCAACGGTCTAGAAATGATTGCTGCAACAACA  2045

Qy  2041 TACTTAAGGATCACAAATTGGATTAGTAGCTGAACATCTCGGATTACCAAACCAACAACCATTC  2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2046 TACTTAAGGATCACAAATTGGATTAGTAGCTGAACATCTCGGATTACCAAACCAACAACCATTC  2105

Qy  2101 GGAACAACTCACCAAGTTCTATCATCCGGTGATCGAGTTCTCATTCAATTAAGAAT  2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2106 GGAACAACTCACCAAGTTCTATCATCCGGTGATCGAGTTCTCATCAATTAAGAAT  2165

Qy  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCAAAATGAAAATAAT  2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2166 ATTAGAACAATCTTCCCTGGTAACCAGTTATGATGATACTTCTTATTCAAAATGAAAATAAT  2225

Qy  2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACCTCAAGTAACCCTGATGCGTCAAGCGCT  2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2226 AAATCTAGTGTGTTTATACATTAAGATTAGCTGACCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2285

Qy  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATTCTTACCTTTATTAGAC  2340
```

Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

```
REFERENCE   1
AUTHORS     Okuda,T., Saito,S., Dorsey,K.M. and Tsuzaki,Y.
TITLE       Modified dna molecule, recombinant containing the same thing, and
            uses thereof
JOURNAL     Patent: EP 1275716-A 79 15-JAN-2003;
            Zeon Corporation (JP)
FEATURES             Location/Qualifiers
     source          1..3129
                     /organism="Mycoplasma galisepticum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:2096"
                     /note="Modified mgc3 gene"
ORIGIN

Query Match             95.9%;  Score 3057.8;  DB 6;  Length 3129;
Best Local Similarity   98.7%;  Pred. No. 0;
Matches 3083;  Conservative  0;  Mismatches  42;  Indels  0;  Gaps  0;

Qy       65 TTGGTTCTGCAAGCTTTGGCTTTAAGCAATCCAGATAGAGAGTAACGATAACACGCAATTAG      124
            |||||||||||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db        5 TCGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAGAGAGTAACGATAACACGCAATTAG       64

Qy      125 TTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGAGACTTGCAGGTCTTGGACAAAATG      184
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       65 TTAATCAAGCAAGAACGCTAGATGCTAATTCGTTGAGACTTGCAGGTCTTGGACAACAGG      124

Qy      185 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGTCTAATG      244
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      125 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGTCTCAGG      184

Qy      245 GAACAATTATCAAATTAGATAGTTTTACTAACAACAAATAGTTTCAGATTACCAACTAGCA      304
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      185 GAACAATTATCAAATTAGATAGTTTTACTAACAACAAATAGTTTCAGATTACCAACTAGCA      244

Qy      305 ATTGTGGTGGATAACAAGTAAAACAAATAGTTTCAGATTACGTCGTTGGTTAATGATGAT      364
            ||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      245 ATTTGCTGGATAACAAGTAAAACAAATAGTTTCAGATTACGTCGTTGGTTAATGATGAT      304

Qy      365 TTGATCAAGAACAAACAAGAGAGCATATATGCTGGTAATTATGCTCGTCGTGTTGGTTAATGATGAAGCTAACGTTC      424
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      305 TTGATCAAGAACAAACAAGAGAGCATATATGCTGGTAATTATGCTCGTCGTGGTTAATGATGAAGCTAACGTTC      364

Qy      425 ATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAATTGGTAATAGAAAACAACAATTCTAAGT      484
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      365 ATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAATTGGTAATAGAAAACCAAAATTCTAAGT      424

Qy      485 TTGTAATTGGTGGTGTTGATAATCCAGTCACGTAACCTGCTGCTGTAAACCTGCT      544
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      425 TTGTAATTGGTGGTGTTGATAATCCAGTCACGTAGTGGTTACTGATGGGACTA      484

Qy      545 AATTAATTTACAACCAAACTCAAGGTGAAATTGTAATCAATCCTTCACTTTCATTTTAGATGCGC      604
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      485 AATTCAGTTACACAGCAAACTCAAGGTGAAATTGTAATGCGAAATTCATTTCAGATGCGC      544

Qy      605 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATATCATTCAAAGAAAGATCT      664
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      545 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATATCATTCAAAGAAAGATCT      604

Qy      665 TACCAAATGACGTCAACACTGCGAGTTGTTCCTGGCCAGTAGGTTAGTGGGGAACAA      724
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      665 TACCAAATGACGTCAACACTGCGAGTTGTTCCTGGCCAGTAGGTTAGTGGGGAACAA      664

Qy      725 ATGCTGATGATGGGATGTTGATTGTGGCGAATGGTCAAATAACTAATAACAGATCCTATTG      784
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      665 ATGCTGATGATGGGATGTTGATTGTGGCGAATGGTCAAATAACTAATAACAGATCCTATTG      724

Qy      785 CTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTCAGGAGCAATGC      844
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      725 CTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC      784

Qy      845 CTGTTGCAAACAATGAGATACGATTCTCAATTGAAGTGCAAGCATAGAATTAAAACATCTT      904
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2286 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGAC     2345
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     2341 AATTCATTCATACAGTAAATGCTGCTGGTAAATGCATTGTTGCTCATCAAACCCTGT     2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2346 AATTCATTCATACAGTAAATGCTGCTGGTAAATGCATTGTTGCTCATCAAACCCTGT     2405

Qy     2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT     2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2406 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT     2465

Qy     2461 GAAGGTTCTGCTGCTAAGTATACATCTGATTTCTGGGGACAATCAAATTCAAACCCGAT     2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2466 GAAGGTTCTGCTGCTAAGTATACATCTGATTTCTGGGGACAATCAAATTCAAACCCGAT     2525

Qy     2521 GAGTACTTAATTCAAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA     2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2526 GAGTACTTAATTCAAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA     2585

Qy     2581 AGCTTCCTTAAACAGTTAGTTGACTTCACTCCGTCTGTAATGCTGCTACTAACTACCGTGTA     2640
            ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     2586 AGCTTCCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGCTACTAACTACCGTGTA     2645

Qy     2641 GTGGTTGATCCTGATGGGTAATTAACAAACAAAACCAAAACCTCAAAAGTTCAGATCCAA     2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2646 GTGGTTGATCCTGATGGGTAATTAACAAACAAAACCAAAACCTCAAAAGTTCAGATCCAA     2705

Qy     2701 TACTTAGATGGTAAGTATTATGATGCTAAATAAAGAACAATAATAATTAGTAACATTCTCT     2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2706 TACTTAGATGGTAAGTATTATGATGCTAAATAAAGAACAATAATAATTAGTAACATTCTCT     2765

Qy     2761 TATAACAACTTGGGGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA     2820
            |||||||||||  ||||||||||||||| |||||||||||||||||||||||||||||||
Db     2766 TATAACAACTTGCTGCTTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTACATTA     2825

Qy     2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCGGTTATCCTTAAGAGCTCAA     2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2826 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCGGTTATCCTTAAGAGCTCAA     2885

Qy     2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT     2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2886 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT     2945

Qy     2941 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTT     3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2946 GCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTT     3005

Qy     3001 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGTAAACCTGCT     3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3006 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACCTGCTGCTGTAAACCTGCT     3065

Qy     3061 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA     3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3066 GCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA     3125

Qy     3121 TCTGGTGCGCCTACAAAACCAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC     3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3126 TCTGGTGCGCCTACAAAACCAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC     3185

Qy     3181 AAAGAATAA     3189
            |||||||||
Db     3186 AAAGAATAA     3194
```

```
Db  785  CTGGTGCAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT  844

Qy  905  TCCAATTAGATGAAAAATTGTTTATCCAGAATGGACTGGTTCTGAGAGAATAAAAATA  964
              ||||||||||||||||| ||||||||||||||||||||||||||||||| |||||||
Db  845  TCCAATTAGATGATGAAAAATTTGTTTATCTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAACAGA  904

Qy  965  TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC  1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  905  TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC  964

Qy  1025 CCGGGACTTCCACAAGTTACTTAAAAGAAGATTCAGTAACGTATTTTCAAGACTATACT  1084
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  965  CCGGGACTTCCACAAGTTACTTTAAAAGAAGATTCAGTAACGTATTTTCAAGACTATACT  1024

Qy  1085 TAAACTCAGTAATTCTTTATCATTCATTGGTGATGATATTTATATTTTTGGTACCTCTG  1144
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1025 TAAACTCAGTAATTCTTTATCATTCATTGGTGATAGTAGTATTTATATTTTTGGTACCTCTG  1084

Qy  1145 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA  1204
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1085 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA  1144

Qy  1205 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATACGGTACAACAACAAACG  1264
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  1145 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATCAGGGTACAACAACACAGG  1204

Qy  1265 GAACAACGACAACAAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGAGAAATACTA  1324
              |||||||||||||||||||||||||||||||||||||||||||||  |||||
Db  1205 GAACAACGACAACAACTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGACAGACTA  1264

Qy  1325 CTAACACTTCTCAAACAGTTTCTAATCCTAAAATACTTATCGTAGTTTTGGAATTG  1384
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1265 CTCAGACTTCTCAAACAGTTTCTAATCCTACTTTAAAATACTTATCGTAGTTTTGGAATTG  1324

Qy  1385 ATAGTAAACCAACTTCTGCAAACAAAAATAGATGAAAATAATTGGCAGATCCTAACGTTA  1444
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  1325 ATAGTAAACCAACTTCTGCAAACAAACAAAATAGATGAAAATAATTGGCAGCAGATCCTAACGTTA  1384

Qy  1445 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA  1504
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db  1385 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAATGAAAATTCCAATAACTA  1444

Qy  1505 ATGCAGGAAACTTTATCCGAAACAAACAAATTGGTGGTGTTACTTCACAAGGTTCAA  1564
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  1445 ATGCAGGAAACTTTATCCGAAACAAACAAATTGGTGGTGTTACTTCACAAGGTTCAA  1504

Qy  1565 GAGTAGTTTTTAAGAGCTTCTAGAAGCTTCCAACTGGAAACTTCCAACTT  1624
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1505 GAGTAGTTTTTAAGAGCTTCTAGAAGCTTCCAACTGGAAACTTCCAACCTT  1564

Qy  1625 TCTTATACGTATTTGGTTATTTAGGATACCAACAACAAGACAACAGGAACTTCTGGTACG  1684
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1565 TCTTATACGTATTTGGTTATTTAGGATACCAACAACAAGACAACAGGAACTTCTGGTACG  1624

Qy  1685 GAACATATAAGCTTTAAACAACAGCCCTATCAGACTGTATTAGATTCTCCAAGAGTAGGTA  1744
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1625 GAACATATAAGCTTTTAAACAACAGCCCTATCAGACTGTATTAGATTCTGCAAGAGTAGGTA  1684

Qy  1745 CTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATTAACTCCATATATAAGAGCACACAC  1804
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1685 CTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATTAACTCCATATATAAGAGCACACAC  1744

Qy  1805 CTGAAGAAGCCGGAAGCATCTTCCAAATCTGGCTATTCTGATAATAATACTTATGAGTACATTCAAT  1864
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1745 CTGAAGAAGCCGGAAGCATCTTCCAAATCTGGCTATTCTGATAATAATACTTATGAGTACATTCAAT  1804

Qy  1865 CAGAAAGCCCAGAGCATCTTCGATAATACTTATGATAATACTTATATGATGAGTACATTCAAT  1924
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1805 CAGAAAGCCCAGAGCATCTTCGATAATACTTATGATAATACTTATATGATGAGTACATTCAAT  1864

Qy  1925 CAGTTTTTAGGATTTGATGGAATTGAATGGAATTTAGAGAAATGTTGGGGTTAAAGCATCAAGCT  1984
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1865 CAGTTTTTAGGATTTGATGGAATTGAATGGAATTTAGAGAAATGTTGGGGTTAAAGCATCAAGCT  1924
```

```
Qy  1985 TCTTAAACTCAAATAGACCCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACATACT  2044
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1925 TCTTAAACTCAAATAGACCCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACATACT  1984

Qy  2045 TAAGATCACAAATTGGATTAGCTAGAGACATCTGGATTACCAAACCAACACCATTCGGAA  2104
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1985 TAAGATCACAAATTGGATTAGCTAGAGACATCTGGATTACCAAACCAACACCATTCGGAA  2044

Qy  2105 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCAGTTCTCATCATCAATTAAGAATATTA  2164
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2045 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCAGTTCTCATCATCAATTAAGAATATTA  2104

Qy  2165 GAACAATCTTCCCTGGTAACCAGTTATAGGTACTTCTTATTCCAAATGAAAATAATAAAT  2224
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATCAGAAAT  2164

Qy  2225 CTAGTGTTTATACATTAAGATATAGCTGACTCAAGTAACCCTGATGCGGTCAAGCTCATTCA  2284
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2165 CTAGTGTTTATACATTAAGATATAGCTGACTCAAGTAACCCTGATGCGGTCAAGCTCATTCA  2224

Qy  2285 GTCCAACAAGTTTAATTGACGTTAATCAAAATTGGTGAATCTTACCTTTATTAGACAATT  2344
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2225 GTCCAACAAGTTTAATTGACGTTAATCAAAATTGGTGAATCTTACCTTTATTAGACAATT  2284

Qy  2345 CATTCTATACAGTAAATGCTGCGTGGTAATGTGTCATTGTTCTCATCAACCCTGGTTCTC  2404
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2285 CATTCTATACAGTAAATGCTGCGTGGTAATGTGTCATTGTTCTCATCAACCCTGGTTCTC  2344

Qy  2405 CTGGATCATATACTGCGTGTAAATACATTTAATCAGAACTTATCGTGATATTGCTTTTGAAG  2464
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2345 CTGGATCATATACTGCGTGTAAATACATTTAATCAGAACTGTGTCGTGATATTGCTTTTGAAG  2404

Qy  2465 GTTCTGGTGCTAAGTAGTATATCATCTGATTTCGGGGAACAATCAATTCAAACCCGATGAGT  2524
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2405 GTTCTGGTGCTAAGTAGTATATCATCTGATTTCTGGGGAACAATCAATTCAAACCCGATGAGT  2464

Qy  2525 ACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAAAGCT  2584
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2465 ACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACACAGCAAAGCT  2524

Qy  2585 TCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTAATAACTACCGGTAGTGG  2644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2525 TCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTAATAACTACCGTGTAGTGG  2584

Qy  2645 TTGATCCTGATGGTAATTTAACAACAAAACCTACCCTAAAAGTTCAGATCCAATACT  2704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2585 TTGATCCTGATGGTCAGCTGACAAAAAACCAAAAACTACCTCTAAAAGTTCAGATCCAATACT  2644

Qy  2705 TAGATGGTAAGTATTATGATGCTAAATTAAGAAACAATAATTAGTAACATTCTCTTATA  2764
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2645 TAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTAGTAACATTCTCTTATA  2704

Qy  2765 ACAACTTTGGCGCCTACCTTACCTTACGTGATAAGAACCTGCTACCAATTGGTAGTAGGTA  2824
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2705 ACAACTTTGGCGCCTACCTTACCTTCTGTGATAAGAACCTGCTACCAATTGGTACTAGGTA  2764

Qy  2825 TTCTTGCAATTATGATCATCTTAGGATTAGGGCATATCATCAAAAAGTTGATACCTTGACTGCTGCTG  2884
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2765 TTCTTGCAATTATGATCATCTTAGGATTAGGGCATATCGTATTCTTAGGAGCTCAAAGAA  2824

Qy  2885 AATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTG  2944
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2825 AATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTG  2884

Qy  2945 TTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAAGTTAAGAAAAACCTGCTG  3004
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2885 TTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAAGTTAAGAAAAACCTGCTCGCTC  2944

Qy  3005 CTTTAGGTGCTGCTGTAAATCTGGTGATAAGGTAAAGCATCAAGCTC  3064
              |||||||||||||||||||||||||||||||||||||||||||||
Db  2945 CTTTAGGTGCTGCTGTAAATCTGGTGATAAGGTAAAGCATCAAGCTC  3004
```

```
Matches 1145;  Conservative  4;  Mismatches  81;  Indels  6;  Gaps  4;

Qy    3065 CAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACCAACTGGGCCTAAATCTG 3124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3005 CAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACCAACTGCGCCTAAATCTG 3064

Qy    3125 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGTTGCTCCAAAACCAACCGCTCCCAAAG 3184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3065 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGTTGCTCCAAAACCAACCGCTCCCAAAG 3124

Qy    3185 AATAA 3189
           |||||
Db    3125 AATAA 3129

RESULT 7
AF083976
LOCUS       AF083976        7141 bp    DNA     linear    BCT 22-AUG-1998
DEFINITION  Mycoplasma gallisepticum adherence protein gene, complete cds.
ACCESSION   AF083976
VERSION     AF083976.1  GI:3450890
KEYWORDS    .
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 7141)
  AUTHORS   Goh,M.S., Gorton,T.S., Forsyth,M.H., Troy,K.E. and Geary,S.J.
  TITLE     Molecular and Biochemical Analysis of a 105 kDa Mycoplasma
            gallisepticum cytadhesin
  JOURNAL   Microbiology (1998) In press
REFERENCE   2  (bases 1 to 7141)
  AUTHORS   Goh,M.S. and Geary,S.J.
  TITLE     Mycoplasma gallisepticum adherence protein gene (gapA) and flanking
            regions
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 7141)
  AUTHORS   Goh,M.S. and Geary,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1998) Pathobiology, University of Connecticut, 61
            North Eagleville Road, Storrs, CT 06269, USA
FEATURES             Location/Qualifiers
     source          1..7141
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="S6"
                     /db_xref="taxon:2096"
                     2994..5888
                     /note="GapA; Mycoplasma pneumoniae P1 homolog; gapA;
                     formerly referred to as MGP1"
                     /codon_start=1
                     /transl_table=4
     CDS             /product="adherence protein"
                     /protein_id="AAC32609.1"
                     /db_xref="GI:3450891"
                     /translation="MQRFNKNKRLKIKCLKTTLCLMHLLVILALKVLQLLLVVVQVV
                     VLLEEVLLLPAASSTARLVEEGNSAGMGTMTPTASTSSTVIDYNSDQNKIPKPKTLLD
                     SSESSESINGGRTYANINTQKNLQGVIVKVNENLFNSENPFAVENMAFIKPKIWVDNY
                     PSTWTQGSANGKWMNVLQFYKHDNPAVNNRFYRAKYYPKRLETQTTTPLIDSSFSPY
                     EHPEWYEGNQFVMPWMQYITNLGGLYAXDGMVYLPGGNGTWVNNESALSIGVFRTKPE
                     NRTAEAPGNTKTVGYPYGILLSAISFDATRNGLALAAPALGQDVGYHFVPRLAVGGVS
                     SPRGANGNIFLPGSSAITWGTNGGNFLDTKWHSPAVIEDAPTTFITVNSSGVLQNSGSQQ
                     STSTPMPNSNGNESIPYRWTNSYDYNSVRFAALISKPAGGNTKQVESLFTTALKLDTL
                     NSLPNKFTQENNIFFSYAMLDGRQWSLGTRKDSTWLTTNTINNPTYNTQQGLASTAAG
                     ENANPRNILNALITAKGFDRRDIGNVVTYSNNTNKFYYYQVGGAITTWPEVQVNYK
                     TSANITYNLTRTDFGSTTPATQDANTVSSKLNGAYLSSTGDQGWYNGSIVKKASF
                     TPSSQGYTWQDFKGLTTTASNAVISNWTKAGYSIRPDDDTVFSVSKIPFEKEITAAVN
                     VRSLDSYYVQLNGETSVNTVARVSPDSSALTLNFKRITNPLMNRDNVIGQGAFISRND
                     IPSSFFENKINDIVTTEADGTEVLDSKYINSIRYTPFPQNNPHIRLRLLVIDRSRATN
                     DFIKLLPQVLVDGEYVAVPQANSVFVSDQEFTGFDALPGYVLPVAISIPIIIALALA
                     LGLGIGIPMSQNRKMLKQGFAISNKKVDILTTAVGSVFKQIINRTSVTNIKKTPQMLQ
                     ANKDGASPSKPSAPAAKKPTGPTKPSAPGAKPTAPAKPKAPATKKIE"

ORIGIN
```

Alignment (Qy/Db pairs):

```
Qy       1 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATACAATTGATAGGTGATTAGCTGTATTTGGA 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5910 ATGAATATTTCTAAAAAACTTAAAAGTATACATTGATACATTGATAGGTGATTAGCTGTATTTGGA 5969

Qy      61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGATAACGATAACACGCAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5970 ACTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGATAACGATAACACGCAA 6029

Qy     121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAAATTCTGTTAGACTTGCGAGGTCTTGGACAA 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6030 TTAGTTAATCAAGCAAGAACGCTAGATGCTAAATTCTGTTAGACTTGCGAGGTCTTGGACAA 6089

Qy     181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6090 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 6149

Qy     241 AATGGACAATTATCAAATTACAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6150 AATGGAACAATTATCAAATTACAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT 6209

Qy     301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCACAACTAGCAGAAAT 360
           |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    6210 GATGATTTTGCTGGTACAAAGTAAAACAAATAGTTTCAGATTACCAACTAGCAGAAAT 6269

Qy     361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6270 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC 6329

Qy     421 GTTCATTCATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACACAATTCT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6330 GTTCATTCATTTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACACAATTCT 6389

Qy     481 AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6390 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG 6449

Qy     541 ACTAAATTAATTTACAAACCAAACTCAAGTGAAATTGTTAA-TGACTTCATTTTAGA 599
           |||||||||||||||||||||| : |||||||||||||||  ||||||||||||||
Db    6450 ACTAAATTTAATTTACAAGCAAACYCAAGGTGAAATTGTTAACCGACTTTATTTTAGA 6509

Qy     600 TGGCGCCAATCTTACCTAAAGATTTACACCCAGATTGGATAATCATTCAAAGAAA 659
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6510 TGGCGCCAATCTTMCCTAAAGATTTACACCCAGATTGGATAATCATTATACATTCAAAGAAA 6569

Qy     660 GATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGG 719
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6570 GATCTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGG 6629

Qy     720 AACAAATGCT-GATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATC 778
           |||||||||| : |||||||||||||  ||:|||||||  |||||:|||:|||||||
Db    6630 TAATAGCGCTACAGGDTGGGACATTTGATTTGATTTCGGGAACGGGMATGCCTTCTGCTAATGATC 6689

Qy     779 CTATTGCTCAAACTCAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAG 838
           ||||||||||||||||||||||||||||||||||    : :  || :|||||||||||
Db    6690 CAATAGCAACGTCTAAAAGCGGTTCTCAAAAAT---AMCCCCTCAAGTTTCAATTCAAGCG 6746

Qy     839 CAATGCCTGGTGCAACAATAGATACGAATTCTCAATTGAATGTCAAGCATAGAATTAAAA 898
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6747 CTCAACCTCGGACAAATAATAGATACGAATTCTCGAAGTTAATGTAAAACACAGAATTAAAA 6806

Qy     899 CATCTTTCCAATTAGATGAAAAATTTGTTATCCAGATGGACTGGTTCTGAAGAGATA 958
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6807 CATCGTTCAATTAGACGAAAAAATTTGTTATCCAGATGAACTGGTTCTGAAGAGAATA 6866

Qy     959 AAAATATTACAAGATTAGCTACTGGAAGTTGCCAGCAACGAAAGATATTGGATTCTTG 1018
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6867 AAATATTACAAGATTACGTACTGGAAGTTGCCAACAACGAAGGAGATATTCAGTTAATTCTG 6926

Qy    1019 ACATACCCGGGACTCCACAAGTACTTTAAAAGAAGATTCAGTTAACGTAATTTCAAGAC 1078
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    6927 ACATACCAGGGACTCCAGAAGTACTTTAAAAGAAGATTCAGTTAACGTAATTTCAAGAC 6986
```

Query Match   33.0%;   Score 1052.4;   DB 1;   Length 7141;
Best Local Similarity 92.6%; Pred. No. 1.2e-167;

```
Qy   121 TTAGTTAATCAAGCAAGAAGACGCTAGATGCTAATTCTGTTGCAGTAGTATTATATTTTGGACAA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTAGTTAATCAAGCAAGAAGACGCTAGATGCTAATTCTGTTGCAGTAGTATTATATTTTGGACAA 180

Qy   181 AATGGTTCGTTGTTCAATCACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATGGTTCGTTGTTCAATCACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240

Qy   241 AATGAACAATTATCATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AATGGAACAATTATCATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGT 300

Qy   301 GATGATTGTGCTGGATGACAAAGTAAAACAAAGTTTCAGATTACAACTAGACAGAAAT 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATGATTTGCTGGATGACAAAGTAAAACAAAGTTTCAGATTACAACTAGACAGAAAT 360

Qy   361 AGATTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC 420

Qy   421 GTTCATTAAAAAGAATTAATACTAACTCAAATAGAATGGTAATAGAAACAACAATTCT 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATGGTAATAGAAACAACAATTCT 480

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATCCAGCTCACGTAATTACTGATGATGGG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG 540

Qy   541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAACGACTTTATTTTAGAT 600

Qy   601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTGTAACTTATACATTCAAAGAAAG 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGATACAACTTATACATTCAAAGAAAG 660

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGGA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGGT 720

Qy   721 ACAAATGCTGATGATGGGGATGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AATAGGCTACAGATGGGGACATTTGATTCGGGAACGGCGATGCCTTCTGCTAATGATCCA 780

Qy   781 ATTGCTCAAACTACAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCA 840
             ||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATAGCAACGTCTAAAAGCGGTTCTAAAAT--AACCCCTCAAGTTCAATTCAATTCAAGCGCT 837

Qy   841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATACAATTAAAACA 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   838 CAACCTGGGACAAACAAAATAATAGATCTGAAGTTAATGTAAAACACAGCATAAATAAACA 897

Qy   901 TCTTTCCAAATGATGAAAAAATTGTTTATCCAGAATGGACTCGGTTCTGAAGAGAATAAA 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 TCGTTTCAATTGACAGAAAAAATTGTTTATCCAGAATGAATGAACTGGTTCTGAAGAGAATAAACA 957

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACAACACGAGATAACAACACGCAA 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAACACAGGAAGATAACAACACGCAA 1017

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1018 ATACCAGGGACTCCAGAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA 1077

Qy  1081 TACTTAAACTCAGTTAATTCTTATCATTCGGATTAGCTGTATTGGA 1134
             ||||||||||||||||||||||||||||||||||||||
Db  1078 TACTTAAACTCAGTTAATTCTTATCATTCATTGGGATTAGCTGTATTGGA 1131
```

RESULT 9
AB033211
LOCUS   AB033211   1128 bp   DNA   linear   BCT 07-APR-2000

```
Qy  1079 TATACTTAAAACTCAGTTAATTCTTATCATCATTGGTGATAGTATTATATTTTTGGGTA 1138
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6987 TATACTTAAAACTCAGTTAATTCTTATCATCATTGGTGATAGTATTATATTTTTGGGTA 7046

Qy  1139 CCTCTGAATTACCATCATTATGGTACTATCATTCCCAACTAGATAATCTGATCATCAACCG 1198
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  7047 CCTCTGAATTACCATCATTATGATACATTATCATTCCCAACTAGATAATCTGATCATCAACCG 7106

Qy  1199 CTTTGAATCAAGTTAAAACAGATGATATTGAAGCTT 1234
             |||||||||||||||||||||||||||||||||||
Db  7107 CTTCTGAATCAAGTT-AAACAGATGATATTCAAGCTT 7141
```

RESULT 8
AB033210
LOCUS   AB033210   1131 bp   DNA   linear   BCT 07-APR-2000
DEFINITION   Mycoplasma gallisepticum gene for 120-kDa membrane protein MGC3,
             partial cds, strain:S6.
ACCESSION   AB033210
VERSION   AB033210.1 GI:7527370
KEYWORDS   120-kDa membrane protein MGC3.
SOURCE   Mycoplasma gallisepticum
ORGANISM   Mycoplasma gallisepticum
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1 (bases 1 to 1131)
AUTHORS   Yoshida,S.
TITLE   Mycoplasma gallisepticum S6-strain gene encoding a 120-kDa membrane
             protein
JOURNAL   Published Only in DataBase (2000)
REFERENCE   2 (bases 1 to 1131)
AUTHORS   Yoshida,S.
TITLE   Direct Submission
JOURNAL   Submitted (04-OCT-1999) Shigeto Yoshida, Jichi Medical School,
             Department of Medical Zoology; Yakushiji 3311-1,
             Minamikawachimachi, Tochigi 329-0498, Japan
             (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339,
             Fax:81-285-44-6489)
FEATURES   Location/Qualifiers
     source   1..1131
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="S6"
                     /db_xref="taxon:2096"
     gene   1..1131
                     /gene="mgc3"
     CDS   1..>1131
                     /gene="mgc3"
                     /codon_start=1
                     /transl_table=4
                     /product="120-kDa membrane protein MGC3"
                     /protein_id="BAA94277.1"
                     /db_xref="GI:7527371"
                     /translation="MNISKLKSYTLIGGLAVFGTLGSASFGFKQSDKSNDNTQLVNQ
                     ARTLDANSVRLAGLGQNGSLFNTVLRDVDDNFITAANGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNRFDQRQTRAYYALLVNDEANVHLKRINTNSNRIGNRNNNS
                     KPVIGGVDNPAHVIRTDDGTKFNFFKQTGEIVNDPILDAPILPKDLHPDWYNLYIQ
                     RKILPNDVNTAVVPWPVGRVSGNSATDGTFDFGNGMPSANDPIATSKSGSQNNPSSFN
                     SSAQPGTNNRYDSEVNVKHRIKTSFQLDEKFVYPEWTGSSENKNITRLATGSLPNNEG
                     YWILDIPGTPEVTLKEDSVAVFSRLYLNSVNSLSFTGDSIYIF"
```

ORIGIN

Query Match          31.1%;    Score 993.2;    DB 1;    Length 1131;
Best Local Similarity 92.9%;    Pred. No. 1.7e-157;
Matches 1053;    Conservative    0;    Mismatches    78;    Indels    3;    Gaps    1;

```
Qy     1 ATGAATATTTCTAAAAAAACTAAAAGTTATACATTGATAAAGTGGATTGGATTAGCTGTATTGGA 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAATATTTCTAAAAAAACTAAAAGTTATACATTGATAAAGTGGATTGGATTAGCTGTATTGGA 60

Qy    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAAGATTCAGATAAGAGTAACGATAACACGCAA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ACTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAAGATTCAGATAAGAGTAACGATAACACGCAA 120
```

```
Qy   361 AGATTTGATCAAAGACAAACAAGAGACATATTATGCTCTGTTGGTAATGATGAAGCTAAC 420
             ||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db   361 AGATTTGATCAAAGACAAACAAGAGACATATTATGCTTTGTTGTAGTTAGTAATGATGTCTATC 420

Qy   421 GTTCATTTAAAAGAAGATTAATACTAACTCAAATAGAATTGGTAATGAAACAACAATTCT 480
             ||||||||| |||||||||||||||||||||||| |||||||||||| |||| ||| ||
Db   421 GTTCATTTAAAAGAAGATTAATAATAACCAAATAGAATTGGTAATTTAAATAAATAATTCG 480

Qy   481 AAGTTGTAATGGTGGTGTGATAATCCAGTCACGTAATTAGATTTACTGATGATGGG 540
         || ||||||||| |  | |||||||||||||||||||||||||||||||||||||||
Db   481 AAATTTGTAATCGGCGTGTTGATAATCCAGTCACGTAATTAGATTTACTGATGATGGG 540

Qy   541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   541 ACTAAATTAATTTAATTTACAAAGCAAACTCAAGGTGAAATTGTTAACGACTTTATTTTAGAT 600

Qy   601 GGGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATCAATTATATCATTCAAAGAAAG 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATCAATTATATCATTCAAAGAAAG 660

Qy   661 ATCTTACCAAATGACGTCAACACTGCAGTGTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA 720
         ||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTAAACACTGCAGTGTGTTCCTTGACCGTGTAGGTAGAGTTAGTGGT 720

Qy   721 ACAAATGCTGATGATGGGATGTTGATTGTGCGGAATGGTCAAATAACTAATAACACAGATCCT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAATGCTACTGATGGGACATTTGATTGTGTAATGGTGAGCTTTTAGTGTAACAAACCCT 780

Qy   781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCA 840
         |||||||||||||| |||| |||||||||||||||||||||||||||||||||||||||
Db   781 ATTGAACAAACTAAAACCACTACCAACAATCAATTGTCTATGACTTTTGATTCATCAGCA 840

Qy   841 ATGCCTGGTGCAAACAATAGATACGATTCCAATTGAATGTCAAGCATAGAATTAAAACA 900
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATGCCTAGTGCAAATATAGATATTAATATAGATTCTGAATTAAATGTCAAACACAGAATTAAAACA 900

Qy   901 TCTTTCCAATTAGATAGAAAAATTGTTTATCCGAATGACAATGGGTTCTGAAGAGAATAAA 960
         ||||| ||||||||||||| |||| |||||||||||||||||||||||||||||||||||||
Db   901 TCGTTTGGATTAGACGAAAAATTTGTTTATCCGAATGACAATGGGTCAAGAAGAGAATGCA 960

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCCTTGAC 1020
         ||||||||||||||||||||||||||||   ------CTAGTGGTAACAATTATTGAATACTTAAT 1014
Db   961 AATCTTACAAGATTAACAAAATAGA------CTAGTGGTAACAATTATTGAATACTTAAT 1014

Qy  1021 ATACCCGGACTCCACAAGTTACTTTAAAAGAAGAGATTCAGTTAACGTAACGTATTTCAAGACTA 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1015 ATTCCTGGAACTCCAAATGTTACTTTAAAAGAACAGACCAGTTAACGTATTTTCAAGATTG 1074

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTTATATTTTT 1134
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1075 TACTTAAACTCAGTTAACTCTTAACTCTCTATCTTTCATCGGTGATAGTATTATTATATTTTT 1128
```

```
RESULT 10
MYCMGP
LOCUS       MYCMGP          8760 bp      DNA      linear      BCT 26-APR-1993
DEFINITION  M.genitalium attachment protein (MgPa) gene, complete cds.
ACCESSION   M31431
VERSION     M31431.1  GI:150157
KEYWORDS    attachment protein.
SOURCE      Mycoplasma genitalium
  ORGANISM  Mycoplasma genitalium
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 8760)
  AUTHORS   Inamine,J.M., Loechel,S., Collier,A.M., Barile,M.F. and Hu,P.C.
  TITLE     Nucleotide sequence of the MgPa (mgp) operon of Mycoplasma
            genitalium and comparison to the P1 (mpp) operon of Mycoplasma
            pneumoniae
  JOURNAL   Gene 82 (2), 259-267 (1989)
  MEDLINE   90060815
   PUBMED   2583522
COMMENT     Original source text: M.genitalium (strain G-37, ATCC 33530) DNA.
FEATURES             Location/Qualifiers
```

```
DEFINITION  Mycoplasma gallisepticum gene for 120-kDa membrane protein MGC3,
            partial cds, strain:F.
ACCESSION   AB033211
VERSION     AB033211.1  GI:7527372
KEYWORDS    120-kDa membrane protein MGC3.
SOURCE      Mycoplasma gallisepticum
  ORGANISM  Mycoplasma gallisepticum
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   Yoshida,S.
  TITLE     Mycoplasma gallisepticum F-strain gene encoding a 120-kDa membrane
            protein
  JOURNAL   Published Only in DataBase (2000)
REFERENCE   2  (bases 1 to 1128)
  AUTHORS   Yoshida,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-1999) Shigeto Yoshida, Jichi Medical School,
            Department of Medical Zoology; Yakushiji 3311-1,
            Minamikawachimachi, Tochigi 329-0498, Japan
            (E-mail:shigeto@jichi.ac.jp, Tel:81-285-58-7339,
            Fax:81-285-44-6489)
FEATURES             Location/Qualifiers
     source          1..1128
                     /organism="Mycoplasma gallisepticum"
                     /mol_type="genomic DNA"
                     /strain="F"
                     /db_xref="taxon:2096"
     gene            1..1128
                     /gene="mgc3"
     CDS             1..>1128
                     /gene="mgc3"
                     /codon_start=1
                     /transl_table=4
                     /product="120-kDa membrane protein MGC3"
                     /protein_id="BAA94278.1"
                     /db_xref="GI:7527373"
                     /translation="MNISKKLKSYTLIGGLAVLGTLGSASFGFKQSDKSNDNTQLVNQ
                     ARTLDANSVKLAGLGQNGSLFNTVLRDVDDNPITAASGTIIKLDSFTKPLYGLDLSDD
                     FAGYKVKQIVSDYTTSRNRFDQRQTRAYYALLVNDDAIVHLKRINTNQNRIGNLNNNS
                     KFVIGGVNDPAHVIRFTDDGTRFKNFTKQTQGEIVNDFILDAPILPKDLHPDWYNLYIQ
                     RKILPNDVNTAVVPWPVGRVSGTNATDGTFDFGNGVAFSTNPIEQTKTTTNNQLSMTF
                     DSSAMPSANNRXDSELNVKHRIKTSFGLDEKKFVYPEWTGQRENANLTRLITNRPSGNNY
                     WILNIPGTPNVTLKEDTVNVFSRLYLNSVNSLASFIGDSIYIF"
ORIGIN
```

```
Query Match     29.0%;  Score 924.8;  DB 1;  Length 1128;
Best Local Similarity  89.2%;  Pred. No. 5.5e-146;
Matches 1011;  Conservative  0;  Mismatches 117;  Indels  6;  Gaps  1;
```

```
Qy    1 ATGAATAATTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAATAATTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTAGGA 60

Qy   61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAAGGTTTAAACAATCAGATAAGAGTAACGATAACACCGCAA 120
        ||||||||||||||||||||||||||||||   |||||| |||| |||||||||||||||||
Db   61 ACTCTTGGTTCTGCAAGCTTTGGCTTTAAACAATCAGATAAGAGTAACGATAACACCGCAA 120

Qy  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTAGACTTCGAGGTCTTGGACAA 180
        |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTAAACTTCGCAGGTCTTGGACAA 180

Qy  181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGAATCAGTAACTGTTAGCTGTTATTAGGA 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AATGGTTCATTGTTCAATACAGTTCTTAGAGATGTTGATGACAACTTTATAACAGCAGCT 240

Qy  241 AATGGAACAATTATCAAATTACAAATTAGATAGTTTTACTAAACAAATAGTTTCAGATCTAAGT 300
        |||||||||||||||||||||||||||| |||||||| |||||||||||||||||||||||
Db  241 AGTGGAACAATTATCAAATTACAAATTAGATAGTTTTTCAGATTACCAAGCCATTAGATCTAAGC 300

Qy  301 GATGATTGTGTGGTGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
        ||| |||| ||||||||||||||||||||||| |||||| |||||||||||||||||||||
Db  301 GATGATTTTGCTGATGATATAAAGTGAAACAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
```

```
source          1..8760
                /organism="Mycoplasma genitalium"
                /mol_type="genomic DNA"
                /db_xref="taxon:2097"
CDS             298..1059
                /note="29 kD protein"
                /codon_start=1
                /transl_table=4
                /protein_id="AAA25419.1"
                /db_xref="GI:150158"
                /translation="MTFTNQIDESFLKEALAIVVDANYKRIRELRELLDKNLFKAVLR
                IDHHPNEDDLNTSFNFVEESYVACCEQIVEMATVAKWTIPPVAATLLYIGIYTDSNRF
                LYSNTSYRTLYLAAILYKAKADIRIVHDHLNHTSLADLKPKKVYYNHFKTQGQVIYFI
                CTKKIQKRLRMTADQCARVNLLSNIADYKIWLFPIEQANNEIRIELRSNGINVRDIAI
                KYGGGGHNNASGAIITNKKQISDVVSDCVRKIVYN"
CDS             1066..5400
                /note="attachment protein MgPa"
                /codon_start=1
                /transl_table=4
                /protein_id="AAA25420.1"
                /db_xref="GI:150159"
                /translation="MHQPKKRLAKKSWAFLTAALTLGVITGVGGYFLFNQNKQRSSVS
                NFAYOPKQLSVKHQQAVDETLIPWTWNNNNPSSLKITGENPGSFGLVRSQNDNLNISS
                VTKNSSDDNLKYLNAVEKYLDGQQNFAIRRYDNNGRALYDINLAKMENFSTVQRGLNG
                EPIFDPFKGFGLTGNAPTDWNEIKGKPVPVVVQSPHSPNLYFVLLVPKVALEYHNLNN
                QVVKESLEVKATQSSFNPTQRLQKDSPVKDSSKQSEKLSETTASSMSSGMATSTRAKA
                LKVEVERGSQSDSLLKNDFAKKPLKHKNSSGEVKLEAEKEFTEAWKPLLTTDQIAREK
                GMGATVVSFYDAPYSENHTAFGLVDHIDPKKMVENYPPSWKTPKWNHHGIWDYNARNL
                LLQTTGFFNPRRHPEWFDEGQAKADNTSPGPKVGDTDHKDGFKQNSSSPIALPFEAY
                FANIGNMVAIGNSVFIFGGNGHATKMFTTNPLSIGVFRIKYTDHSKSSVTGWFYAVL
                FGGLINPQTNGLKDLPLGTNRWFEVYPRMAVSGVNGQLVLAGTLTMGDTAVTRL
                KYDQLEKHLANLVAQQGLLREDLQIPTPYGWANRPDIPVGAWLQDEMGSKFGPHYFLN
                NPDIQDNVNNDTVEALISSYKNTDKLKHVYPYRYSGLYAWQLFWSNKLLNTPLSANF
                VNENSVAPNSLFAAILNEDLLTGLSDKIFYGKENEFAENEADRFNQLLSLNPNNTNW
                ARYLNVVQRFTTGPNLDSSTPDQFLDPLEWIGNGKPFSNSPSPSTSASSTPLPTFSN
                INVGVKSMITOHLNKENTRWVFIPNFSPDIWTGAGYRVQSANQKNGIPFEQVKPSNNS
                TPFDPNSDDNKVTPSCGSSKPTTYPALPNSISPTSDWINALTFTNKNNPQRNQLLLRS
                LLGTIPVLINKSGDSNDQFNKDSEQKWDKTETNEGNLPGFGEVNGLYNAALLHTYGFP
                GTNTNSTDPKIGFKADSSSSSSSSTLVGSGLNWTSQDVGNLVVINDTSFGFQLGGWFIT
                FTDFIRPRTGYLGITLSSLQDQTIIWADQPWTSFKGSYLDSDGTPKSLWDPTALKSLP
                NSSTTYDTNPTLSPSFQLYQPNKVKAYQTTNTYNKLIBPVDATSAATNWTSLLKLLTT
                KNIKAKLGKTPASSQGNNNGGVSQQTINITTTTGNISEGLKEETSIQAETLKKFPDSK
                QNNKSEIGIGDSTFFKMDGKLTGVVSTPLVNLINGQGATSDSDTEKISFKPGNQIDFN
                RLFTLPVTELFDPNTMFVYDQYVPLLLVNLPSGFDQASIRLKVISYSVENQTLGVNGF
                KDPQTQFIPVLNASSTGPQTVFQPFNQWADYVLPLIVTVPIVVIILSVTLGLTIGIP
                MHRNKKALQAGFDLSNKKVDVLTKAVGSVFKEIINRTGISNAPKKLKQATPTKPTKT
                PPKPPVKQ"
CDS             5402..8560
                /note="114 kD protein"
                /codon_start=1
                /transl_table=4
                /protein_id="AAA25421.1"
                /db_xref="GI:150160"
                /translation="MKTMRKQIYKKAYWLLLPFLPLALANTFLVKEDSKNVTAYTPFA
                TPITDSKSDLVSLAQLDSSYQIADQTIHNTNLPVLFKSRDVKVKYESSGSNNISFDST
                SQGEKPSYVVEFTNSTNIGIKWTMVCKYQLDVPNVSSDMNQVLKNLLLEQLTKYTLN
                SSLAKEKGKTQREVHLGSGQANQWTSQRNQHDLNNNPSPNASTGFKLTTGNAYRKLSE
                SWPIYEPIDGTKQGKGKDSGGWSSTENEAKNDAPVSGGGSSSGTFNKYLNTKQALE
                SIGILFDDQTPRNVITQLYYASTSKLAVTNNHIVMGNSFLPSMWYWVVERSAQENAS
                NKPTWFANTNLDWGEDKQKQFVENQLGYKETTSTNSNHFHSKSFTQPAYLISGIDSVN
                DQIIFSGFKAGSVGYDSSSSSSSSTKQOALAWSTTTSLDSKITGYKDLVTNDFGLN
                GPINGSFSIQDTFSFVVPYSGNHTKNGTTGPIKTAYPVKDQKSTVKINSLINATPLN
                SYGDEGIGVFDALGLNYNFKSNQERLPSRTDQIFVYGIVSPNELRSAKSADSTGSDT
                KVNWSNTQSRYLPVPYNYSEGIIDADGFKRPENRGASVTTFSGLKSIAPDGFANSIAN
                FSVGLKAGIDPNPVMSGKKANYGAVVLIKHVPNVSSDMNQVLKNLLSTTDNNIAPISFS
                FTPFTAAESAVDLTTFKEVTYNQESGLWSYIFDSSLKPSHDGKQTPVTDNMGFSVITV
                SRTGIELNQDQATTLDVAPSALAVQSGIQSTTQTIFGVLPLSEEFSAVIAKDSDQNK
                IDIYKNNGLFEIDTQLSNSVATNNGGLAPSYTENRVDAWGKVEFADNSVLQARNLVD
                KTVDEIINTPEIINSFRFTPAFEDKATLVATKQSDTSLSVSPIQFLDGNFYDLNS
                TIAGVPLNIGFPSRVFAGFAALPAWVIVSVGSSVGILFILLVLGLGIGIPMYRVRKL
                QDASFVNVFKKVDTLTTAVGSVYKKIITQTGVKKAPSALKAANPSVKKPAAFLKPPV
                QPPSKPFGEQKAVEVKSEETKS"
misc_signal     8582..8613
```

```
RESULT 11
U39698
LOCUS       U39698         15787 bp    DNA     linear     BCT 05-NOV-1998
DEFINITION  Mycoplasma genitalium section 20 of 51 of the complete genome.
ACCESSION   U39698 L43967
VERSION     U39698.1  GI:3844782
KEYWORDS    .
SOURCE      Mycoplasma genitalium
  ORGANISM  Mycoplasma genitalium
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 15787)
  AUTHORS   Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,
            Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.G.,
            Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M.,
            Fuhrmann,J.L., Nguyen,D.T., Utterback,T.R., Saudek,D.M.,
            Phillips,C.A., Merrick,J.M., Tomb,J.-F., Dougherty,B.A., Bott,K.F.,
            Hu,P.C., Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.
  TITLE     The minimal gene complement of Mycoplasma genitalium
  JOURNAL   Science 270 (5235), 397-403 (1995)
  MEDLINE   96026346
  PUBMED    7569993
REFERENCE   2  (bases 1 to 15787)
  AUTHORS   Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,
            Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.,
            Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V.,
            Sandusky,M., Fuhrmann,J.L., Nguyen,D.T., Utterback,T.R.,
            Saudek,D.M., Phillips,C.A., Merrick,J.M., Tomb,J.-F.,
            Dougherty,B.A., Bott,K.F., Hu,P.-C., Lucier,T.S., Peterson,S.N.,
            Smith,H.O., Hutchinson,C.A. III and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-OCT-1995) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 15787)
  AUTHORS   Fraser,C.M., Gocayne,J.D., White,O., Adams,M.D., Clayton,R.A.,
            Fleischmann,R.D., Bult,C.J., Kerlavage,A.R., Sutton,G.G.,
            Kelley,J.M., Fritchman,J.L., Weidman,J.F., Small,K.V., Sandusky,M.,
            Fuhrmann,J.L., Nguyen,D.T., Utterback,T.R., Saudek,D.M.,
            Phillips,C.A., Merrick,J.M., Tomb,J.-F., Dougherty,B.A., Bott,K.F.,
            Hu,P.C., Lucier,T.S., Peterson,S.N., Smith,H.O. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-1998) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
COMMENT     On Nov 5, 1998 this sequence version replaced gi:1045891.
FEATURES             Location/Qualifiers
     source          1..15787
```

```
                /organism="Mycoplasma genitalium"
                /mol_type="genomic DNA"
                /isolate="G37"
                /db_xref="taxon:2097"
gene            1264..2016
                /gene="MG186"
CDS             1264..2016
                /gene="MG186"
                /note="similar to GB:U00089 SP:P75265 PID:1673668 percent
                identity: 44.35; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=4
                /product="lipoprotein, putative"
                /protein_id="AAC71405.1"
                /db_xref="GI:3844783"
                /translation="MKGFLKPNFSLGALFLTLSPIATACIAEKPVNNRFNFNSEQLAR
                LRKARVNHWRDGDTLSVSFANNHQKPIRIYAIDTPEKAVLSIQRKSEIELKEANKATE
                FAKSLIPIGSEVWIWPINSYSYDREVAAVFFKTNPLQLHFSFAVEMVANGHALPIAG
                NDPDFVFSDLDPFNPLKIVGIELANGLNNAFNNRKNIFSYLENSFQSITMVYQQRGVD
                QSWTRYLAPSNDFSSTKLGLGLTIYELKLNNG"
gene            2009..3766
                /gene="MG187"
CDS             2009..3766
                /gene="MG187"
                /note="similar to GB:U00089 SP:P75264 PID:1673667 percent
                identity: 61.74; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=4
                /product="ABC transporter, ATP-binding protein"
                /protein_id="AAC71406.1"
                /db_xref="GI:3844784"
                /translation="MANNKSAIELKNIVVDFGESVAIDNINLSVEKGHQLVSLLGPSGC
                GKTTTLAVIAGLIKPTSGQVLFNGYDVTKKPPQERKIGLVFQNYALYPHMNVFENIVP
                PLYSDNSWKQAVLEKNSVANHEINCLLLTSNGASVQEIDQLNKLFHDSIEKPKQIQYQ
                INDLNVSVFKNLNELTANLKLIPSKHQFAITNLNKQTLKQINELEAEFKTKWKLQKQT
                PIKSGVEHNAKLQAIKQHFSYEKQRLKKGHYFKTKVELKQTIVENLKLVKKAISEQTKL
                IKQSSDYTKLQLKRLLKIKVEPNQLKKQYKVPLNQLIKNYSLKTDKLTDTQLNEIEQIK
                TRIVSIKQFINKTALEVANKLAITKLLTKPDKISGGQQQRVAIARAIVRRPKLLLMD
                EPLSNLDAKLRQTRQWIRQFQQBLQITTVFVTHDQBEAMSISDVIVCMSTGKVQQIG
                TPSELYLKPANEFVARFIGTFPEMNIIECSVKNQLFWNNHLLVTESFKLNVEKLLVGF
                RYEQLVVTTNKSSLQAKLINIENLGKHLVATISLFDTTLSWRLELNSHLKVGDSLNFI
                IKANNLHFFDITKQRIEI"
gene            3769..4758
                /gene="MG188"
CDS             3769..4758
                /gene="MG188"
                /note="similar to GB:U00089 SP:P75263 PID:1673666 percent
                identity: 74.09; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=4
                /product="ABC transporter, permease protein"
                /protein_id="AAC71407.1"
                /db_xref="GI:3844785"
                /translation="MFKWLLKHNQPHSLQLGLLDQPLPFWKPFLLPALLTTLLFT
                IPFFLSLQKGFSANSDLYDLSSQSFSLRTFQDLFESNFVLGLRNSFLYSLISILPFS
                IIIAIVIASAIVFVYKKLLRGFWQTVFLIPYVTSGVAISIAFVYIFDSASGILNTVFN
                VNTKWLDSGSRDTFNALWAILIFGVWKNLAFNVLIISTAMLSVNPGLYKVASLDSANP
                VRQFFKITLPSIRPTLIFLTLTLLILGCMQVFPLALFENKPEEAVANGGNSILLYIFQQ
                IQSGNTNLAGAATLVLFVLGVCYGLVLRNGFVLIEWLQWKKIKQLYVQKQLTLY"
gene            4682..5686
                /gene="MG189"
CDS             4682..5686
                /gene="MG189"
                /note="similar to GB:U00089 SP:P75262 PID:1673665 percent
                identity: 69.67; identified by sequence similarity;
                putative"
                /codon_start=1
                /transl_table=4
                /product="ABC transporter, permease protein"
                /protein_id="AAC71408.1"
                /db_xref="GI:3844786"
                /translation="MAFIWLNGCSEKLNSFMFKNNLRFTSWINQHKFYQLDLSLKTRS
                IKQIVLTLVFKTLVLGFPGLIVIPPFYLMVVVSFASDERALDTRTPILWPDSWNFDNF
                SRVLSDGKYLNAIVNTLVTVLSVLLTLFFTICMGYSFSLRKWKYKKLVWFFLSVLI
                LPESALLIGQYRIVIVAWNNFNPSPLIVLGLIMPFVSSVFSGFMYRTSFSAIFSQLKB
                SALIDGCNGFNYFLKIALPMVKSTSWTVGILTAFSAWNSYLWPLLLGNRVDLINLW
                VLQQGILDANSSDEQIRTLNLKMSAAILAILPMFIIYPLFHKRIMNAIKNRANTIKG
gene            5726..6811
                /gene="MG190"
CDS             5726..6811
                /gene="MG190"
                /note="similar to PID:1673660 percent identity: 60.38; identified by
                sequence similarity; putative"
                /codon_start=1
                /transl_table=4
                /product="mgpA protein (mgpA)"
                /protein_id="AAC71409.1"
                /db_xref="GI:3844787"
                /translation="MGKTSTISNHCVIRDYLLSVCHLWHCASDHYHYQWRGTFIFVM
                KKGSTEAINAIKQPDKIVIFHHVRPDGDCLGAQQGLFLHLIKANFKNKEVKCVGNNNN
                LFSPINMTFNQIDESPLKEALAIVVDANYKNRIELRELLDKNLPKAVLRIDHHPNED
                DLNTSFNFVBESYVACCEQIVEMATVAKWTIPPVAATLLYIGIYTDSNRFLYSNTSYR
                TLYLAAILYKAKADIRIVHDHLNHTSLADLKFKKCVYNHFKTCQQVIYPICTKKIQKR
                LRMTADQCARVNLLSNIADYKIWLFPIEQANNEIRIDLRSNGINVRDIAIKYGGGHEN
                NASGAIITNKKQISDVVSDCVKKIVYN"
gene            6818..11152
                /gene="MG191"
CDS             6818..11152
                /gene="MG191"
                /note="similar to SP:P20796 GB:M31431 GB:U01694 GB:U01779
                GB:U02159 percent identity: 98.89; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=4
                /product="adhesin P1 (mgpB)"
                /protein_id="AAC71410.1"
                /db_xref="GI:3844788"
                /translation="MHQPKRLAKKSWAFLTAALTLGVITGVGYFLFNQNKQRSSVS
                NFAYQPKQLSVKHQQAVDETLTPWTWNNNNFSSLKITGENPGSFGLVRSGNDLNISS
                VTKNSSDDNLKYLNAVEKYLDGQQNFAIRYDNGNRALYDINLAKMENPSTVQRGLNG
                EPIFDPFKGFGLTGNAPTDWNEIKGKVPEVVQSPHSPNLYFVLLVPKVALEYHNLNN
                QVVKESLEVKATQSSFNPTQRLQKDSPVIDQSLKQBKLSETTASSMSSGMATSTRAKA
                LKVEVERGSQSDSLLKNDFAKKPLKHKNSSGBVKLEAEKEFTEAWKPLLTTDQIAREK
                GMGATVVSFYDAPYSENHTAFGLVDHIDPKKMVENYPPSWKTPKWNHHGIWDYNARNL
                LLQTTGFFNPRRHPEWPDEGQAKADNTSPGFKWQTDHKKDGPFKQNSSSPIALPPEAY
                FANIGNMVAIGNSVPIFPGNGHATKMFTNPLSIGVPRIKYTDNPSKSSVTGWPYAVL
                FGGLINPQTNGLKDLPLGTNRWFEYVPRMAVSGVKWVGNQLVLAGTLIMGDTATVPRL
                KYDQLEKHLNLVAQQGLLREDLQIFTPYGWANRPDIPVGAWLQDEMGSKFGPHYFLN
                NPDIQDNVANDTVEALISSYKNTDKLKHVYPYRSGLYAWQLFNWSNKLTNTPLSANF
                VNENSYAPNSLFAAILNEDLLTGLSDKIFYGKENBFAENEADRFNQLLSLNPNNTNW
                ARYLNVVQRFTTGPNLDSSTFDQFLDFLPWINGKPFSNSPSPSTSASSSTPLPFSN
                INVGVKSMTQHLKENTRWVFIPNFSPDIWTGAGYRVQSANQKNGIPPEQVKPSNNS
                TPFDPNSDDNKVTPSGGSSKPTTVPALPNSISPTSDWINALTFTNKNNPQRNQLLRS
                LLGTIPVLINKSGDSNDQFNKDSEQKWDKTETNEGNLPGFGEVNGLYNAALLHTYGFF
                GTNTNSTDPKIGFKADSSSSSSSTIVGSGLNWTSQDVGNLVVINDTSFGQLGGWFIT
                FTDFIRPRTCYLGTTLSSLQDQTIIWADQPWTSPKGSYLDSDGTPKLSWDPTALKSLP
                NSSTTYDTNPTLSPSPQLYQPNKVKAYQTTNTYNKLIEPVDATSAATNMTSLLKLLTT
                QNIKAKLAGTLASSCGNNNGGGVSQTINTTITTGNISEGLKEETSIQAETLKKFFDSK
                QNNKSBIGIGDSTFTKMDGLTGVVSTPLLNLINGGQGATSDSTBKISFKPGNQIDFSN
                RLFTLPVTELFDPNTMFVYDQYVPLLVNLPSGFDQASIRLKVISYSVENQTLGVRLBP
                KDPQTQQFIPVLNASSTGPQTVFQPFNQWADYVLPLIVVPIVVIILSVTLGLTIGIP
                MHRNKKALQAGFDLSNKKVDVLTKAVGSVFKEIINRTGISNAPKLKQATPTKPTPKT
                PPKPPVKQ"
gene            11154..14312
                /gene="MG192"
CDS             11154..14312
                /gene="MG192"
                /note="similar to SP:P22747 GB:U02157 GB:M31431 GB:U02161
                GB:X61522 percent identity: 99.43; identified by sequence
                similarity; putative"
                /codon_start=1
```

```
RESULT 13
MYCATTP
LOCUS        MYCATTP        9691 bp    DNA    linear    BCT 15-FEB-1996
DEFINITION   M.pneumoniae P1 operon encoding P1 surface protein, ORF-4, and
             ORF-6 proteins, complete cds.
ACCESSION    M21519 M20916
VERSION      M21519.1  GI:150138
KEYWORDS     P1 attachment protein; P1 surface protein.
SOURCE       Mycoplasma pneumoniae
  ORGANISM   Mycoplasma pneumoniae
             Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE    1  (bases 1054 to 5937)
  AUTHORS    Inamine,J.M., Denny,T.P., Loechel,S., Schaper,U., Huang,C.H.,
             Bott,K.F. and Hu,P.C.
  TITLE      Nucleotide sequence of the P1 attachment-protein gene of Mycoplasma
             pneumoniae
  JOURNAL    Gene 64 (2), 217-229 (1988)
  MEDLINE    88297153
  PUBMED     2841195
REFERENCE    2  (bases 1 to 9691)
  AUTHORS    Inamine,J.M., Loechel,S. and Hu,P.C.
  TITLE      Analysis of the nucleotide sequence of the P1 operon of Mycoplasma
             pneumoniae
  JOURNAL    Gene 73 (1), 175-183 (1988)
  MEDLINE    89211947
  PUBMED     2468577
COMMENT      Original source text: Mycoplasma pneumoniae (strain M129) (tissue
             library: ATCC 29342) DNA.
             Draft entry and computer-readable sequence for [2] kindly provided
             by S.Loechel, 16-NOV-1988.
FEATURES             Location/Qualifiers
     source          1..9691
                     /organism="Mycoplasma pneumoniae"
                     /mol_type="genomic DNA"
                     /strain="M129"
                     /db_xref="taxon:2104"
                     /tissue_lib="ATCC 29342"
     mRNA            54..9616
                     /product="P1 operon mRNA"
     CDS             316..1041
                     /note="ORF-4; putative"
                     /codon_start=1
                     /transl_table=4
                     /product="unknown protein"
                     /protein_id="AAA88324.1"
                     /db_xref="GI:1196657"
                     /translation="MQQALAVIVDANYKERIECRDLLDQNQFKAVLRIDHHPNEDDLN
                     TTHNFVDASYIAAAEQVVDLAVQAKWKLSPPAATALYLGIYTDSNRFLYSNTSWRTLY
                     LGSMLYRAQANIAKIHDELNHTSLKDIQFKQVYFKNFQTFQNVIYFVADKKFQKKLKV
                     TPLECARVNILANIEQFHIWLFFIEBGKMHYRVBFRSNGINVREVALKYGGGGHIQAS
                     GAVLKSKRDIIRVVQDCQKQIAV"
     CDS             1054..5937
                     /codon_start=1
                     /transl_table=4
                     /product="P1 protein"
                     /protein_id="AAA88325.1"
                     /db_xref="GI:150139"
                     /translation="MHQTKKTALSKSTWILILTATASLATGLTVVGHFTSTTTLKRQ
                     QFSYTRPDEVALRHTNAINPRLTPWTYRNTSFSSLPLTGENPGAWALVRDNSAKGITA
                     GSGSQQTTYDPTRTEAALTASTTFALRYDLAGRALYDLDFSKLNPQTPTRDQTQGIT
                     FNPPGGFGLGSGAAPQQWNEVKNKVPVEVYAQDPSNPYRPAVLLVPRSVVYVEQLQRGLG
                     GSSGDTKVVALKIBVKKKSDEDNGQLQLEKNDLANAPIKRSEBSGQSVQLKADDFGT
                     ALSSSGSGGNSNPGSPTPWRPWLATEQIHKDLPKWSASILILYDAPYARNRTAIDRVD
                     HLDPKAMTANYPPSWRTPKWNHHGLWDWKARDVLLQTTGFFNPRRHPENFDGGQTVAD
                     NEKTGPDVDNSENTKQGFQKEADSDKSAPIALPPEAYFANIGNLTWFGQALLVPGGNG
                     HVTKSAHTAPLSLGVFRVRVNATGSAATVTGWPYLALLFSGMVNKQTDGLKDLPNNNR
                     MFEYVPRMAVAGAKFVGRELVLAGIITMGDTATVTPRLLYDBLBSNLNLVAQGGLLRE
                     DLQLFTPYGWANRPDLPIGAWSSSSSSHNAPYYFHNNPDWQDRPIQNVVDAFIKPWE
                     DKNGKDAKYIYPYRYSGMWAWQYNWSNKLTDQPLSADFVNENAYQPNSLFAAILNP
                     ELLAALPDKVKYGKENEFAANEYERFNQKLTVAPTQGTNWSHFSPTLSRFSTGFNLVG
```

```
/transl_table=4
                     /product="mgpC protein (mgpC)"
                     /protein_id="AAC71411.1"
                     /db_xref="GI:384789"
                     /translation="MKTNRKQIYKKAYWLLLPFLPLALANTFLVKBDSKNVTAYTPFA
                     TPITDSKSDLVSLAQLDSSYQIADQTIHNTNLFVLFKSRDVKVKYBSSGSNNISFDST
                     SQGEKPSYVVEFTNSTNIGIKWTMVKKYQLDVPNVSSDMNQVLRNLIILSQPLTKYTLN
                     SSLAKEKGKTQREVHLGSQQANQWTSQRNQHDLNNPSPNASTGFKLTTGNAYRKLSE
                     SWPIYEPIDGTKQGKGKDSGWSSTBENEAKNDAPSVSGGGSSGTFNKYLNTKQALE
                     SIGILFDDQTPRNVTQLYYASTSKLAVTNNHIVVMGNSFLPSWYWVVERSAQENAS
                     NKPTWFANTNLDWGEDKQKQPFVENQLGYKETTSTNSHNFHSKSFTQPAYLISGIDSVN
```

```
Query Match       3.9%;   Score 124.6;  DB 1;  Length 15787;
Best Local Similarity  63.5%;  Pred. No. 1.9e-11;
Matches 190;   Conservative  0;   Mismatches 109;   Indels  0;   Gaps  0;

Qy   2770 TTTGGGCGCCTTACCTTCCATGGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  13950 TTTGCAGCACTCCCTGCATGGGGTGATCCCTGTATCGTAGTAGGTTCTTCAGTTGGGATCTTG 14009

Qy   2830 GCAATTATGATCATCTTAGGATTAGCTATCGTATCCTTTAAGAGCTCAAAGAAAATTA 2889
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14010 TTTATCTTGTTAGTCTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 14069

Qy   2890 CAAGACAAAGGGTTCAAAACAACACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14070 CAAGATGCATCGTTTGTTAATGTCTTTTAAAAAGGTTGATACACTCACAACTGCTGTCGGT 14129

Qy   2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTGCTGCTTTA 3009
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14130 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGTGTAAAAAGCACCTAGTGCATTG 14189

Qy   3010 GGTGCTCGTGTAAATCTGGTGATAAGAAAACCTGCTGCTCGCTGTTCAGC 3068
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  14190 AAAGCTGCTAATCCTAGTGTGTAAAAAACCTGCTGCTGTTTTTAAAAACCACCTGTTCAACC 14248
```

```
RESULT 12
AR300198_2
WPCOMMENT
Continuation (3 of 6) of AR300198 from base 200001 (AR300198 Sequence 1 from patent US 6
Sequence split into 6 fragments  LOCUS AR300198 Accession AR300198
 Fragment Name        Begin          End
AR300198_0              1           110000
AR300198_1           100001         210000
AR300198_2           200001         310000
AR300198_3           300001         410000
AR300198_4           400001         510000
AR300198_5           500001         580073

Query Match       3.9%;   Score 124.6;  DB 6;  Length 110000;
Best Local Similarity  63.5%;  Pred. No. 1.2e-11;
Matches 190;   Conservative  0;   Mismatches 109;   Indels  0;   Gaps  0;

Qy   2770 TTTGGGCGCCTTACCTTCCATGGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28703 TTTGCAGCACTCCCTGCATGGGGTGATCCCTGTATCGTAGTAGGTTCTTCAGTTGGGATCTTG 28762

Qy   2830 GCAATTATGATCATCTTAGGATTAGCTATCGTATCCTTTAAGAGCTCAAAGAAAATTA 2889
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28763 TTTATCTTGTTAGTCTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 28822

Qy   2890 CAAGACAAAGGGTTCAAAACAACACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28823 CAAGATGCATCGTTTGTTAATGTCTTTTAAAAAGGTTGATACACTCACAACTGCTGTCGGT 28882

Qy   2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCTGCTGCTTTA 3009
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28883 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGTGTAAAAAGCACCTAGTGCATTG 28942

Qy   3010 GGTGCTCGTGTAAATCTGGTGATAAGAAAACCTGCTGCTCGCTGTTCAGC 3068
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  28943 AAAGCTGCTAATCCTAGTGTGTAAAAAAACCTGCTGCTGTTTTTAAAAACCACCTGTTCAACC 29001
```

```
Db         9534  GCACCAGTTAAACACGTCTGCTCCAACAGCTCCAAGACCACCAGTCCAACCACCTA  9588
                 |||| ||||||| || ||| | ||||  ||||| ||| |||| || |||||||||||
```

```
RESULT 14
AE000002
LOCUS       AE000002              16876 bp    DNA     linear   BCT 24-NOV-2000
DEFINITION  Mycoplasma pneumoniae M129 section 15 of 63 of the complete genome.
ACCESSION   AE000002 U00089
VERSION     AE000002.2  GI:11379441
KEYWORDS    .
SOURCE      Mycoplasma pneumoniae
  ORGANISM  Mycoplasma pneumoniae
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 16876)
  AUTHORS   Himmelreich,R., Hilbert,H., Plagens,H., Pirkl,E., Li,B.C. and
            Herrmann,R.
  TITLE     Complete sequence analysis of the genome of the bacterium
            Mycoplasma pneumoniae
  JOURNAL   Nucleic Acids Res. 24 (22), 4420-4449 (1996)
  MEDLINE   97105885
   PUBMED   8948633
REFERENCE   2  (bases 1 to 16876)
  AUTHORS   Dandekar,T., Huynen,M., Regula,J.T., Ueberle,B., Zimmermann,C.U.,
            Andrade,M.A., Doerks,T., Sanchez-Pulido,L., Snel,B., Suyama,M.,
            Yuan,Y.P., Herrmann,R. and Bork.P.
  TITLE     Re-annotating the Mycoplasma pneumoniae genome sequence: adding
            value, function and reading frames
  JOURNAL   Nucleic Acids Res. 28 (17), 3278-3288 (2000)
  MEDLINE   20411492
   PUBMED   10954595
REFERENCE   3  (bases 1 to 16876)
  AUTHORS   Himmelreich,R., Hilbert,H. and Li,B.-C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1996) Zentrun fuer Molekulare Biologie
            Heidelberg, University Heidelberg, 69120 Heidelberg, Germany
REFERENCE   4  (bases 1 to 16876)
  AUTHORS   Suyama,M., Dandekar,T. and Herrmann,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2000) Zentrun fuer Molekulare Biologie
            Heidelberg, University Heidelberg, 69120 Heidelberg, Germany
COMMENT     On Nov 27, 2000 this sequence version replaced gi:1673651.
            This updated annotation replaces the old annotation from reference
            1. The old gene identifiers (MP numbers) according to the original
            publication by
            Himmelreich et al. (1996) are given as well as new gene numbering
            (MPN numbers) from the origin of replication. Annotation comments
            and further update data are at
            http://www.bork.embl-heidelberg.de/Annot/MP/.
FEATURES             Location/Qualifiers
     source          1..16876
                     /organism="Mycoplasma pneumoniae"
                     /mol_type="genomic DNA"
                     /strain="M129"
                     /db_xref="taxon:2104"
                     /note="ATCC 29342"
     gene            138..1112
                     /gene="orf4"
                     /note="synonym: B07_orf324"
     CDS             138..1112
                     /gene="orf4"
                     /note="MPN140(new), 015 (Himmelreich et al., 1996)"
                     /codon_start=1
                     /transl_table=4
                     /product="DHH family phosphoesterases"
                     /protein_id="AAB95662.1"
                     /db_xref="GI:1673660"
                     /translation="MNSQVRRKGSIAEAVSAIQAHDKIVIFHHIRPDGDCLGAQHGLA
                     RLIQTNFPHKQVFCVGDPKENFPWLEMKFKQITPELMQQALAVIVDANYKERIEC
                     RDLLDONQFKAVLRIDHHPNEDDLNTTHNFVDASYIAAAEQVDLAVOAKWKLSPPAA
                     TALYLGIYTDSNRPLYSNTSWRTLYLGSMLYRAQANIAKIHDBLNFTSLKDIQFKQYV
                     FKNFQTFQNVIYFVADKKFQKKLKVTPLECARVNILANIEQPHIMLFPIEGKNHYRV
                     EFRSNGINVREVALKYGGGHIQASGAVLKSRDIIRVVQDCKQIAV"
```

```
        SVLDQVLDYVPWIGNGYRYGNNHRGVDDITAPQTSAGSSSGISTNTSGGRSFLPTPSN
        IGVGLKANVQATLGGSQTMITGGSPRRTLDQANLQLWTGAGWRNDKASSGQSDENHTK
        TPTADWPNALSPTNKNNAQRAQLFLRGLLGSIPVLVNRSGSDSNKPGAYTNLPPNL
        TPTADWPNALSPTNKNNAQRAQLFLRGLLGSIPVLVNRSGSDSNKPGAYTNLPPNL
        HSDQTKLNLPAYGEVNGLLNPALVETYFGNTRAGSGSGNTTSSPGIGFKIPEQNNDSK
        ATLIITPGLAWTPQDVGNLVVSGTTVSFQLGGWLVTFTDFVKPRAGYLGLQLTGLDASD
        ATQRALTWAPRPWAAFRGSWNVRLGRVESVWDLKGVWADQAQSDSQGSTTTAIRNALP
        EHPNLAFQVSVVEASAYKPNTSSGQTQSTNSSPYLHLVKPKKVTQSDKLDDDLKNLL
        DPNQVRTKLRQSFGTDHSTQPQPSGLKTTTPVFGTSSGNLSSVLSGGAGGGSSGSGQ
        SGVDLSPVEKVSGWLVGQLPSTSDGNTSSTNLAPVTNTGNDVVGVGRLSBSNAAKMN
        DDVDGIVRTPLAELLDGEGQTADTGPQSVKFKSPDQIDFNRLFTHPVTDLFDPVTMLV
        YDQYIPLFIDIPASVNPKMVRLKVLSPDTNRQSLGLRLEFFKPDDTQPNNNVQVNPN
        NGDFLPLLTASSGQPQFLFSSPNQWPDYVLPLAITVPIVVILVVLSVTLGLAIGIPHHKN
        KQALKAGFALSNQKVDVLTKAVGSVPKEIINRTGISQAPKRLKQTSAAKPGAPRPPVP
        PKPGAPKPVPVQPPKKPA
CDS     5943..9599
                /note="ORF-6; putative"
                /codon_start=1
                /transl_table=4
                /product="P1 protein"
                /protein_id="AAA88326.1"
                /db_xref="GI:1196658"
                /translation="MKSKLKLKRYLLFLPLLPLGTLSLANTYLLQDHNTLLTPYTPFTT
        PLNGGLDVVRAAHLHPEYELVDWKRVGDTKLVALVRSALVRVKFQDTTSDQSNTNQN
        ALSFDTQESQKALNGSGSGSSDTSGSNSQDFASYVLIFKAAPRATWVFERKIKLALPY
        VKQESQQSGDQSNGKGSLYKTLIQDLLVEQPVTPYTPNAGLARVNGVAQDTVHEGSGQ
        ESSWNSQRSQKGLKNNGFPKAVTGFKLDKGRAYRKLNESWPVYEPLDNSSLGSTKEGKDES
        SWKNSEKTTAENDAPLVGMVSSGAAGSASSLQGNGSNSSGLKSLLRSAPVSVPPSSTS
        NQTLSLSNPAFVGPQAVVSQPAGGATAAVSVNRTASDTAIFSKYLNTAQALHQMGVIV
        PGLEKWGGNNGTVVASRQDATSTNLPNAGASQTGLGTGSPREPALTATSQRAVTVV
        AGPLRAGNSSETDALPNVITQLYHTSTAQLAYLNGQIVVMGSDRVPSLWTWVVGEDQE
        SGKATWWAKTELNWGTDKQKQFVENQLFKODDSNDSKNSNLKAQGLTQPAYLIRGLD
        VVADHLVFAAFKAGAVGYDMTTDSSASTYNQALAWSTTAGLDSDGGYKALVENTAGLN
        GPINGLPTLLDTFAYVTPVSGMKGGSQNNEEVQTTYPVKSDQKATAKIASLINASPLN
        SYGDDGVTVFDALGLNFNFKLNEERLPSRTDOLLVYGIVNESELKSARENAQSTSDDN
        SNTKVKWTNTASHYLPVPYYYSANFPEAGNRRAEQRNGVKISTLESQATDGFANSLL
        NFGTGLKAGVDPAPVARGHKPNYSAVLLVRGGVVRLNFNPDTDKLLDSTDKNSEPISF
        SYTPFGSABSAVDLTTLKDVTYIAESGLWFYTFDNCGEKPYTDGKQQVNRKGYAVIT
        VSRTGIEFNEDANTTTLSQAPAALAVQNGIASSQDDLTGILPLSDEFSAVITKQQTWT
        GKVDIYKNTNGLFEKDDQLSENVKRRDNGLVPIYNEGIVDIWGRVDFAANSVLQARNL
        TDKTVDEVINNPDILQSFPKFPTPAPDNQRAMLVGEKTSDTTLTVKPKIEYLDGNFYGE
        DSKIAGIPLNIDFPSRIFAGFAALPSWVIPVSVGSSVGILLILLLLGLGIGIPMYKVR
        KLQDSSFVDVFKKVDTLTTAVGSVYKKIITQTSVIKKAPSALKAANNAAPKAPVKPAA
        PTAPRPPVQPPKKA"
   terminator   9617..9651
                /note="putative"
   ORIGIN  526 bp upstream of HindIII site.

Query Match             3.2%;  Score 102.2; DB 1;  Length 9691;
Best Local Similarity   55.5%;  Pred. No. 1.3e-07;
Matches  197;  Conservative   0;  Mismatches  158;  Indels   0;  Gaps   0;

Qy   2737  AACAATAAATTAGTAACATTCTCTATAACAACTTGGCGCCTTACCTTCATGGGTAGTG  2796
           |||| |||||   ||| || |||| || |||||| |||| ||| || ||||| ||||
Db   9234  AACATTGATTTCCCTTCCCGGATTTTGCTGGCTTTTACCGTCCTCGGTCATT  9293

Qy   2797  CCTACAGCAATTGGTAGTACATTAGGTATTCTTGCAATTATGATCATCTTAGGATTAGCT  2856
           ||||  ||  |||| |||||| |||||||  || || ||| |||| |||| ||||||
Db   9294  CCGGTATCAGTCGGTTCATCGGTGGGCATTCTCTTAATCCTGCTCATCTTAGGCCTTGGT  9353

Qy   2857  ATCGGTATTCCTTTAAGAGCTCAAAGAAAATTACAAGACAAAGGGTTCAAAACAACATTC  2916
           || ||| |||    ||||  | |||||  ||  |||  |||| ||||| ||| || ||||
Db   9354  ATTGGAATTCCAATGTATAAAGGTCGCAAGCTTCAAGACTTCAAGCTTTGTTGATGTGTTT  9413

Qy   2917  AAAAAAGTTGATACCTTGACTGCGTGCTGTGGTTCAGTTACAAGAAGATTATTACCCCAA  2976
           |||||  |||| ||  | |||||| || || ||||| || ||||||||| || |||||
Db   9414  AAAAAGGTGGATACGTTGACAACCCGCTGTGGGTAGCGTGTACAAGAAGATTATCACCCCA  9473

Qy   2977  ACTGCTAACGTTAAGAAGAAAAAAACCTGCTGCTTTAGGTGGTGGTAAATCTGGTGATAAGAAA  3036
           || |||| || ||||| |||||| ||| ||||  |||||||| || ||||| ||||| |||
Db   9474  ACGAGTGTGATCAAAAAGTCCTAGTCCGTTGAAAGCTGCTAATAACGCTGCTCCTAA  9533

Qy   3037  CCTGCTGCTAACCTGCTCGCTCCAGCTAAACCATCTGCACCAAAAGCTA  3091
```

```
gene            1125..6008
                /gene="P1"
CDS             1125..6008
                /note="synonym: E07_orf1627"
                /gene="P1"
                /note="MPN141(new), 014(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="ADP1 MYCPN adhesin P1"
                /protein_id="AAB95661.1"
                /db_xref="GI:1673659"
                /translation="MHEQTKKTALSKSTWILILTATASLATGLTVGHFTSTTTLKRQ
                QFSYTRPDEVALRHTNAINPRLTPWTYRNTSFSSLPLTGENFGAWALVRDNSAKGITA
                GSGSQQTTYDPTRTEAALTASTTPALRRYDLAGRALYDLDPSKLNPQTPTRDQTGQIT
                FNPFGGFGLSGAAPQQWNEVKNKVPVBVAQDPSNPYRFAVLLVPRSVVYTEQLQRGLG
                LPQQRTBSGQNTSTTGAMFGLKVKNAEADTAKSNEKLQGAEATGSSTTSGSGQSTQRG
                GSSGDTVKALKIEVKKKSDSEDNGQLQLEKNDLANAPIKGSEBSGQSVQLKADDFGT
                ALSSSGSGSNSNPGSPTPWRPWLATEQIHKDLPKWSASILLIYDAPYARNRTAIDRVD
                HLDPKAMTANYPPSWRTPKWNHHGLWDKARDVLLQTTGFFNPRRHPEWFDGGQTVAD
                NEKTGFDVDNSENTKQGFQKEADSDKSAPIALPPEAYPANICNLTWPGQALLVFGGNG
                HVTKSAHTAPLSIGVFRVRYRNATGTSAIVTGWPYALLFSGMVNKQTDGLKDLPFNNNR
                WFEIVPRMAVAGAKFVGRELVLAGTITMGDTATVPRLLYDELESNLNIVAQCGLLRE
                DLQLFFTPYGWANRPDLFIGAWSSSSSSSSHNAPYFHNNPDWQDRPIQNVVDAFIKPWE
                DKNGKDDAKYIYPYRYSGMWAWQYNWSNKLTDQPLSADFVNENAYQPNSLFAAILNP
                ELLAALPDKVKYGKENEPAANEYBRFNQKLTVAPTQGTNWSHFSPTLSRFSTGFNLVG
                SVLDQVLDYVPWIGNGYRYGNNHRGVDDITAPQTSAGSSSGISTNTSGSRSFLPTFSN
                IGVGLKANVQATLGGSQTMITGGSPRRTLDQANLQLWTGAGWRNDKASSGQSDENHTK
                FTSATGMDQQGQGSTSAGNPDSLKQDNISKSGDSLTTQDGNAIDQQEATNVTNLPPNL
                TFTADWPNALSFTNKNNAQRAQLFLRGLLGSIPVLVNRSGSDSNKFQATDQKWSYTDL
                HSDQTKLNLPAYGEVNGLLNPALVETYFGNTRAGGSGSNTTSPGIGFKIPEQNNDSK
                ATLITPGLAWTPQDVGNLVVSGTTVSFQLGGWLVTFTDFVKPRAGVLGLQLTGLDASD
                ATQRALIKAPRPWAAFRGSWVNRLGRVESVWDLKGVWADQAGSDSQGSTTTRTRNALP
                EHPNALAFQVSVVEASAYXPNTSSGQTOSTNSSPYLHIVKPKKVTGSDKLDDLKNLL
                DPNQVRTKLRQSFGTDHSTQPOPQSLKTTTPVFGTSSGNLSSVLSGGGAGGSSGSGGQ
                SGVDLSPVEKVSGWLVGQLPSISDGNISSTNNLAPNTINTGNDVVGVGRLSESNAAKWN
                DDVDGIVRTPLAELLDGEGQTADTGPQSVKFKSPDQIDFNRLFTHPVTLDFPVTMLV
                NGDFLPLLTASSGQPQT.FSPFNQWPDVLPLAITTVPIVIVLSVTLGLAIGIPMEKN
                KQALKAGFALSNQKVDVLJTKAVGSVFKIIINRTGISQAPKRLKQTSAAKPGAPRPPVP
                PKPGAPKPVQPKKPA"
repeat_region   1125..2671
                /rpt_family="REPMP4"
repeat_region   3171..4971
                /rpt_family="REPMP2/3"
gene            6014..9670
                /gene="E07_orf1218"
                /note="synonym: orf6"
gene            6014..9670
                /gene="orf6"
CDS             6014..9670
                /gene="orf6"
                /note="MPN142(new), 013(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="involved in cytadherence"
                /protein_id="AAB95660.1"
                /db_xref="GI:1673658"
                /translation="MKSKLKLKRYLLFLPLLPLGTLSLANTVLLQDENTLTPYTPFT
                PLNGGLDVVRAAHLHPSYELVDWKRVGDTKLVALVRSALIVRVKFQDTTSSDQSNTNQN
                ALSFDTQESQKALNGSQSGSSDTSGSNSQDFASYVLIPKAAPRATWVFBRKIKLALPY
                VKQESQGSGSDDQSGNGKGSLYKTLQDLLVEQPVTPYTPNAGLARVNGVAQDTVHFGSGQ
                ESSWNSQRSRSQKGLKNNPGPKAVTGFKLDKGRAYRKLNESWPVYEPLDSTKEGKGKDES
                SWKNSEKTAENDAPLVGMVSGAAGSASSLQGNGNSSGLSKLLRSAPVSVPPSSTS
                NQTLSLSNPAPVGPQAAVVSQPAGGATAAVSVNRTASDPATFSKYLNTAQALHQMGVIV
                PGLEKWGGNNGTGVVASRQDATSTNLPHAAGASQTGLGTGSPREPALTATSQRAVTVV
                AGPLRAGNSSETDALPNVITQLYHTSTAQLAYLNGQIVMGSDRVPSLWYWVGEDQE
                SGKATWAKTELNWGTDKQKQFVENQLGFKDDSNSDKSNSNLKAQGLTQPAYLIAGLD
                VVADHLVPAAFKAGAVGVDMTTDSSASTYNQALAWSTTAGLDSDGYKALVENTAGLN
                GPINGLFTLLDTFAYVTPVSGMKGGSQNNEEVQTTYPVKSDQKATAKIASLINASPLN
                SYGDDGTVFDALGLNFNFKLNEFKLKSARENAQSTSDDN
                SNTKVKWTNTASHYLPVPVYYSANFPEAGNRRAEQRNGVKISTLESQATDGFANSLL
                NFGTGLKAGVDPAPVARGHKPNYSAVLLVRGGVVRLNFNPDTDKLLDSTDKNSEPISF
                SYTPFGSABSAVDLITLKDVTYIAESGLWFYTFDNGEKPTYDGKQQVKNRKGYAVIT
```

```
repeat_region   6014..8714
gene            9787..10314
                /rpt_family="REBPMP5"
gene            9787..10314
                /gene="E07_orf175"
CDS             9787..10314
                /gene="E07_orf175"
                /note="MPN143(new), 012(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein"
                /protein_id="AAB95659.1"
                /db_xref="GI:1673657"
                /translation="MSHKDFNGLQAPQLLSSSSPVAKKQSHKLRHALKHARYLNHSS
                KRTLKHALELHEDNQVLLEBEGSPNFQDWLSKQPGVNKTSLKYNKSLGSWISKESKPK
                KRFPPYFTYKGSKTTPEBAKALQQMKQSQKRFFHENMHSFLNEVAHNPMIQRFKQKQA
                KRAANTRQRTYKYRQ"
repeat_region   10457..11977
                /rpt_family="REBMP2/3"
gene            10888..12129
                /gene="E07_orf413"
CDS             10888..12129
                /gene="E07_orf413"
                /note="MPN144(new), 011(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein, see: MPN141"
                /protein_id="AAB95658.1"
                /db_xref="GI:1673656"
                /translation="MCQQGQGSTSAGNPDSLKQDKISKSGDSLTTQDGNATGQQEATN
                YTNLPPNLTPTADWPNALSFTNKNNAHRAQLFLRGLLGSIPVLVNRSGSDSNKFQATD
                QKWSYTDLQSDQTKLNLPAYGEVNGLLNPALVETYFGNTRAGGSGSNTTSPGIGFKI
                PEQNNDSKATLITPGLAWTPQDVGNLVVSGTSLSFQLGGWLVSFTDFIKPRAGVLGLQ
                LSGLDASDSDQRELIWAKRPWAAFRGSWVNRLGRVESVWDLKGVWADQAQLAAQAATS
                BASGSALAPHNALAFQVSVVEASAYSSSTSSGSGSSNTSPYLHLIKPKKVESTTQ
                LDQGLKNLLDPNQVRTKLRQSFGTDHSTQPOSLKTTTPVFGTSSGNIGSVLSGGGAGG
                GSSGSGGSGVDLSPVERVSGH"
gene            12481..13020
                /gene="E07_orf179"
CDS             12481..13020
                /gene="E07_orf179"
                /note="MPN145(new), 010(Himmelreich et al., 1996)"
                /codon_start=1
                /transl_table=4
                /product="hypothetical protein, see: MPN013"
                /protein_id="AAB95657.1"
                /db_xref="GI:1673655"
                /translation="MSYSPSLKIBIIAILQKYTSKNYQSNCKTRPDGKLELSLNGVFBE
                IVKTPGKTRYVTHKQLDQKLKDFKQDLMVELHDTATKTDLKESEARINQKLEALIQI
                VMVQGEQIKVEGBQINKLTQAVEKQGEKIEAQGQQIQKVNETLNFVVESLGSIHKRLD
                SMEGRLDSMENRLDKLBSK"
gene            13338..14135
                /gene="E07_orf265"
CDS             13338..14135
                /gene="E07_orf265"
```

```
VSRTGIEFNEDANTTLSQAPAALAVQNGIASSQDDLTGILPLSDFSAVITKDQTWT
GKVDIYKNTNGLFEKDDQLSSNVKRRDNGLVPIYNEGIVDIWGRVDPAANSVLQARNL
TDKTVDEVINNPDIIQSFFKFTPAFDNQRAMLVGEKTSDTTLTVKPKIEYLDGNFYGE
DSKIAGIPLNIDPPSRIFAGFAALPSWVIPVSVGSSVGILLILLIGLGIGIPMYKVR
KLQDSSFVDVFKKVDTLTTAVGSVYKKIITQTSVIKKAPSALKAANNAAPKAPVKPAA
PTAPRPPVQPPKKA"
```

```
Query Match         3.2%;    Score 102.2;    DB 1;    Length 16876;
Best Local Similarity 55.5%;  Pred. No. 1.1e-07;
Matches 197;  Conservative  0;  Mismatches 158;  Indels  0;  Gaps  0;

Qy   2737  AACATAATTAGTAACATTCTCTTATAACAACTTGGCGCCTTACCTTCATGGGTAGTG  2796
           ||||  |||| |||| ||| |||||| ||||||  || |||||  |||| |||||||||
Db   9305  AACATTGATTTCCCTTCCCGGATTTTGCTGGCTTTGCTTTGCTTTGCTTTACCGTCCTGGGTCATT  9364

Qy   2797  CCTCAGCAATTGGTACATTAGGTATTCTTGCAATTATGATCATCTTAGGATTAGCT  2856
           || ||||  |||| |||||| ||||||  ||  ||   |||||  |||||||| |||
Db   9365  CCGGTATCAGCAGTCGGTTCATCGGTCGGCATTCCTTAATCCTTAATCCTGCTCATCTTAGGCCTTGGT  9424
```

```
Qy   2857  ATCGGTATTCCTTTAAGAGCTCAAAGAAAATTACAAGACAAAGGGTCAAAACAACATTC  2916
                  || ||||| ||||| ||        ||||||      |||||||||
Db   9425  ATTGGAATTCCAATGTATAAGGTCCGCAGCTTCAAGACTCCAGCTTTGTTGTTGATGTGTTT  9484

Qy   2917  AAAAAAGTTGATACCTTGACTGCGTCTGTTGGTTCAGTTACAAGAAGATTATTACCCAA  2976
                  ||||| |||  ||  |||||| |||  | |||||||| |||||||||
Db   9485  AAAAAGGTGGATACGTTGACAACCGTCGTGGGTAGCGTGTACAAGAAGATTATCACCCAA  9544

Qy   2977  ACTGCTAACGTTAAGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAA  3036
                  ||    |  ||| | ||| | || | |||||||||||||||| |||| ||
Db   9545  ACGAGTGTGATCAAAAAAGCTCCTAGTGCGTTGAAAGCTGCTATAACCTGCTCCTAA  9604

Qy   3037  CCTGCTGCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTA  3091
                  |||||     |    ||||  |||  ||| |   |||||  |||
Db   9605  GCACCAGTTAAACCAGCTGCTCCAACAGCTCCAAGACCACCAGTCCAACCACCTA  9659
```

RESULT 15

MGU34970

```
LOCUS       MGU34970     1618 bp   DNA   linear   BCT 22-FEB-1996
DEFINITION  Mycoplasma genitalium repetitive sequence element mgp-r5.
ACCESSION   U34970
VERSION     U34970.1  GI:1177840
KEYWORDS    .
SOURCE      Mycoplasma genitalium
  ORGANISM  Mycoplasma genitalium
            Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE   1  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Hu,P.C., Bott,K.F. and Hutchison,C.A. III.
  TITLE     A survey of the Mycoplasma genitalium genome by using random
            sequencing
  JOURNAL   J. Bacteriol. 175 (24), 7918-7930 (1993)
   MEDLINE  94075230
   PUBMED   8253680
REFERENCE   2  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Bailey,C.C., Jensen,J.S., Borre,M.B., King,E.S.,
            Bott,K.F. and Hutchison,C.A. III.
  TITLE     Characterization of repetitive DNA in the Mycoplasma genitalium
            genome: possible role in the generation of antigenic variation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 92 (25), 11829-11833 (1995)
   MEDLINE  96102208
   PUBMED   8524858
REFERENCE   3  (bases 1 to 1618)
  AUTHORS   Peterson,S.N., Bailey,C.C., King,E.S., Bott,K.F. and Hutchison,C.A.
            III.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-AUG-1995) Clyde Hutchison, Microbiology and
            Immunology, The University of North Carolina, CB# 7290 F.L.O.B.,
            Chapel Hill, NC 27599-7290, USA
FEATURES             Location/Qualifiers
     source          1..1618
                     /organism="Mycoplasma genitalium"
                     /mol_type="genomic DNA"
                     /strain="G37"
                     /db_xref="taxon:2097"
                     /clone="eshl0"
     CDS             <1..270
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); ORF3 protein"
                     /codon_start=1
                     /transl_table=4
                     /protein_id="AAA88899.1"
                     /db_xref="GI:1177841"
                     /translation="IPMRVRKLQDASFVNVFKKVDTLTTAVGSVYKKIITGTGVVKK
                     APSALKAANPSVKKPAAFLKPPVQPPSKPEGEQKAVEVKSEETKS"
     CDS             <658..>1068
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); mgp-r5; similar to B
                     region of MgPa adhesin"
                     /codon_start=1
```

```
                     /transl_table=4
                     /protein_id="AAA88900.1"
                     /db_xref="GI:1177842"
                     /translation="LLVPKVVLEYHKLSKDVVKESLEVEATDSFDPTQRLQKDSPMKD
                     SSKQGEKLSETTASSMSGMATSPRKALKIEVEKGSSGSDTLTKSDFAKKPLKHKENSG
                     TEVKLDAQKDFAGEKAWKPLLTTDQIAREKGMGAT"
     CDS             <1068..>1618
                     /note="submitter believes protein coding sequences are not
                     expressed unless the sequence recombines into the
                     functional adhesin operon (MgPa); mgp-r5; similar to EF
                     region of MgPa adhesin"
                     /codon_start=1
                     /transl_table=4
                     /protein_id="AAA88901.1"
                     /db_xref="GI:1177843"
                     /translation="LDFLPWIGNNKPFSNSHTASLSVSSNTPLFTFSNINVGVKSMIT
                     QHLNKENRWVFTPNSSPDIWTGAGYRKQGNNNGIPFDNVKPSNSSTPFNPNSDDNKVT
                     SGGSSKPTYTHLPNSISPTSDWSNALTFTNKNNPQRNQLLLRSLLGLGLTIPVLINKSGT
                     GDQFNKDSEQKWNETDKLGGNLPG"
ORIGIN
```

```
Query Match          2.9%;  Score 93;  DB 1;  Length 1618;
Best Local Similarity  68.3%;  Pred. No. 6.7e-06;
Matches  129;  Conservative  0;  Mismatches  60;  Indels  0;  Gaps  0;

Qy   2880  AAGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGC  2939
                  |||||||              |||||||  ||||||||||| ||||| |||
Db     18  AAGAAAACTCCAAGATGCATCGTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAAC  77

Qy   2940  TGCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACC  2999
                  ||||| || || | |||| |||||||| |||| |||||  |||| || |||||||
Db     78  TGCTGTCGGTAGTGTGTACAAAAAGATTATTACCCAAACTGTGTGTGAAAAAAGCACC  137

Qy   3000  TGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCTGC  3059
                  |||||||   ||||||  ||| |||||||||||||  ||||||| |||||||
Db    138  TAGTGCATTGAAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGTTTTTAAAACCACC  197

Qy   3060  TGCTCCAGC  3068
                  |||||
Db    198  TGTTCAACC  206
```

Search completed: May 5, 2004, 17:06:54
Job time : 11960.1 secs

OM nucleic - nucleic search, using sw model

```
            GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.
```

Run on: May 5, 2004, 09:54:21 ; Search time 1116.68 Seconds
                                    (without alignments)
                          12131.925 Million cell updates/sec

Title: US-09-901-572A-2
Perfect score: 3189
Sequence: 1 atgaatattctcaaaaaact........caaccgctcccaaagaataa 3189

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
        1: geneseqn19805:*
        2: geneseqn19905:*
        3: geneseqn20005:*
        4: geneseqn2001as:*
        5: geneseqn2001bs:*
        6: geneseqn20025:*
        7: geneseqn2003as:*
        8: geneseqn2003bs:*
        9: geneseqn2003cs:*
       10: geneseqn20045:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3189	100.0	3189	7	ACF03365	Acf03365 Mycoplasm
2	3144.2	98.6	3189	2	AAT75087	Aat75087 DNA encod
3	3143.4	98.6	8354	5	ABX95108	Abx95108 DNA encod
4	3143.4	98.6	10651	7	AAF88025	Aaf88025 Vector pc
5	3057.8	95.9	3129	7	ACF03435	Acf03435 Mycoplasm
6	124.6	3.9	3159	2	ACA39446	Aca39446 Prokaryot
7	124.6	3.9	110000	2	AAT58840_2	Continuation (3 of
8	78.4	2.5	4182	2	AAQ81778	Aaq81778 DNA encod
c 9	64.6	2.0	3435	5	AAQ79746	Aaq79746 Mycoplasm
c 10	62.6	2.0	4985	6	ABQ75107	Abq75107 KSHV LUR
c 11	62.6	2.0	4985	9	ACF79720	Acf79720 Anopheles
c 12	60.8	1.9	8771	3	ABL33825	Abl33825 Human imm
c 13	60.2	1.9	3489	5	AAA30290	Aaa30290 Kaposi's
c 14	60.2	1.9	3489	4	AAF82901	Aaf82901 Nucleotid
15	60.2	1.9	3489	6	ABA93487	Aba93487 Kaposi's
c 16	60.2	1.9	32207	2	AAV73805	Aav73805 KSHV LUR
c 17	60.2	1.9	137507	2	AAV19941	Aav19941 KSHV long
c 18	59	1.9	535	6	ABQ42292	Abq42292 Oligonucl
19	59	1.9	535	6	ABQ42293	Abq42293 Oligonucl
20	58.4	1.8	110000	6	ABA92787_5	Continuation (6 of
c 21	58	1.8	50000	6	ABL55644	Abl55644 AmEPV gen
c 22	55.8	1.7	969	6	ABQ39490	Abq39490 Oligonucl
c 23	55.8	1.7	969	6	ABQ39491	Abq39491 Oligonucl
c 24	55	1.7	2000	7	ADA71938	Ada71938 Rice gene
25	54.8	1.7	3837	3	AAA70211	Aaa70211 Plasmodiu
26	54.6	1.7	14041	4	AAH48024	Aah48024 Internal
27	53.8	1.7	3582	3	AAA70241	Aaa70241 Plasmodiu
28	53.4	1.7	1159	3	AAS59240	Aas59240 An EcoRI
29	53.2	1.7	987	3	AAT90555	Aat90555 Clostridi
30	53.2	1.7	1037	3	AAS59242	Aas59242 Exons E,
31	53.2	1.7	1472	3	AAS59241	Aas59241 Exons D,
32	53.2	1.7	8056	7	ABZ10246	Abz10246 Haematopo
33	53	1.7	987	2	AAT90554	Aat90554 Clostridi
c 34	53	1.6	13784	6	ABK40062	Abk40062 Human che
35	52.4	1.6	3057	2	AAX99555	Aax99555 Nucleic a
36	52	1.6	972	2	AAX61754	Aax61754 B. burgdo
37	52	1.6	1029	4	AAX61753	Aax61753 B. burgdo
c 38	52	1.6	5120	2	AAC84677	Aac84677 DNA seque
c 39	52	1.6	110000	2	AAX20248_08	Continuation (9 of
40	51.8	1.6	8056	7	ABZ10100	Abz10100 Haematopo
41	51.6	1.6	906	7	ACA28937	Aca28937 Prokaryot
42	51.6	1.6	2451	4	AAS07045	Aas07045 DNA encod
43	51.4	1.6	2553	7	ACA52806	Aca52806 Prokaryot
44	50.8	1.6	2017	9	ADB46163	Adb46163 Plasmodiu
45	50.8	1.6	5314	9	ADB46161	Adb46161 Plasmodiu
```

ALIGNMENTS

```
RESULT 1
ACF03365
ID   ACF03365 standard; DNA; 3189 BP.
XX
AC   ACF03365;
XX
DT   11-SEP-2003  (first entry)
XX
DE   Mycoplasma gallisepticum mgc3 gene SEQ ID NO:2.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.
XX
PF   11-JUL-2002; 2002EP-00254879.
XX
PR   11-JUL-2001; 2001US-00301572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG ) ZEON CORP.
XX
PI   Okuda T,  Saito S,  Dorsey KM,  Tsuzaki Y;
XX
DR   WPI; 2003-373746/36.
DR   P-PSDB; ABR57375.
XX
PT   DNA molecule derived from a prokaryotic cell, useful for producing a
PT   vaccine for treating viral infections comprises at least one modified DNA
PT   regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS   Claim 9; Page 31-32; 70pp; English.
XX
CC   The present invention describes a DNA molecule derived from a prokaryotic
CC   cell, where at least one of the DNA regions encoding NXB (where N =
CC   asparagine, X = any amino acid other than proline, and B = serine or
CC   threonine) has been modified so that no N-glycosylation occurs during the
CC   expression in a eukaryotic cell. Also described: (1) a fused DNA
CC   molecule, where a DNA encoding a signal sequence has been ligated to the
CC   N-terminal end of the modified DNA molecule as described above so that it
CC   may be expressed as a fusion protein; (2) a recombinant virus integrated
CC   with the DNA molecule or the fused DNA molecule described above; (3)
```

CC producing a modified or fusion protein by using the recombinant virus
CC described above, to express a protein encoded by the modified DNA
CC molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC vaccine comprising the recombinant virus. The DNA molecule has virucide
CC and immunostimlant activities. The DNA molecule is useful for producing
CC a vaccine for treating viral infections. The present sequence is used in
XX the exemplification of the present invention
SQ Sequence 3189 BP; 1114 A; 576 C; 539 G; 960 T; 0 U; 0 Other;

Query Match   100.0%; Score 3189; DB 7; Length 3189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 ATGAATAATTTCTAAAAAAACTTAAAAGTTATACATGATAGGTGGATTAGCTGTATTTGGA    60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAATAATTTCTAAAAAAACTTAAAAGTTATACATGATAGGTGGATTAGCTGTATTTGGA    60

Qy    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGAGTTAAGTGATAGAGTAACACGCAA   120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 GCTCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGAGTTAAGTGATAGAGTAACACGCAA   120

Qy   121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGCAGACAA   180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGCAGACAA   180

Qy   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT    240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT    240

Qy   241 AATGGAACAAATATCAAATTAGATAGTTTACTAAAACAAATAAGTTGATGGTTTAGATCTAAGT   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AATGGAACAAATATCAAATTAGATAGTTTACTAAAACAAATAAGTTGATGGTTTAGATCTAAGT   300

Qy   301 GATGATTGTGTGGATACAAAGAGCAAAACAAGAGCAATATATGCTCGTGGTTAATGATGAAGCTAAC   360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATGATTGTGTGGATACAAAGAGCAAAACAAGAGCAATATATGCTCGTGGTTAATGATGAAGCTAAC   360

Qy   361 AGATTTGATCAAGACAAAGACAAGAGAGCAATTAGCTCGTGTTGGTAATGATGAAGCTAAC   420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGATTTGATCAAGACAAAGACAAGAGAGCAATTAGCTCGTGTTGGTAATGATGAAGCTAAC   420

Qy   421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGTAATAGAAACAACAATTCT   480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTCATTTAAAAAGAATTAATACTAACTCAAATAGAATTGTAATAGAAACAACAATTCT   480

Qy   481 AAGTTTGTAATTGGTGTGTTGATAATCCAGCTCACGTAATTACTGATTTACTGATGATGGG   540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGTTTGTAATTGGTGTGTTGATAATCCAGCTCACGTAATTACTGATTTACTGATGATGGG   540

Qy   541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT   600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT   600

Qy   601 GCGCCAATCTTACCTAAAGATTACACCCAGTTGTGTTCCTTGGGCTATATCAAAGAAAAG   660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGCCAATCTTACCTAAAGATTACACCCAGTTGTGTTCCTTGGGCTATATCAAAGAAAAG   660

Qy   661 ATCTTACCAAATGACGTCAACACTGGATGTTTGATTGTGGGAATGGTCAAAATCCTTCAACTTTTA   720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTCAACACTGGATGTTTGATTGTGGGAATGGTCAAAATCCTTCAACTTTTA   720

Qy   721 ACAAATGCTGATAAAACCACTAACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA   780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAATGCTGATAAAACCACTAACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA   780

Qy   781 ATTGCTCAAACTAAAAACCACTAACTGATAATCCTTCAACTTTTAATTCAGGAGCA   840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATTGCTCAAACTAAAAACCACTAACTGATAATCCTTCAACTTTTAATTCAGGAGCA   840

Qy   841 ATGCCTGGTGCAAACATAGATACGATTCTCAATTGAATGTCAAGCATAAATAAAACA   900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATGCCTGGTGCAAACATAGATACGATTCTCAATTGAATGTCAAGCATAAATAAAACA   900

Qy   901 TCTTTCCAATTAGATGAATGAAAAATTTGTTATTTCAAGAATGGACTGGTTCTGAAGAGAATAAA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTCCAATTAGATGAATGAAAAATTTGTTATTTCAAGAATGGACTGGTTCTGAAGAGAATAAA   960

Qy   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020

Qy  1021 ATACCCGGACTCCACAAGTTACTTAAAAGAAGAGATTCAGTTAACGTATTTCAAGACTA  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATACCCGGACTCCACAAGTTACTTAAAAGAAGAGATTCAGTTAACGTATTTCAAGACTA  1080

Qy  1081 TACTTAAACTCAGTAATTCTTATCATTCATTGGTGATAGTATTTATATTTTGGTACC  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TACTTAAACTCAGTAATTCTTATCATTCATTGGTGATAGTATTTATATTTTGGTACC  1140

Qy  1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT  1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT  1200

Qy  1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCAGTCTTGCAGGTACAACAACA  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCAGTCTTGCAGGTACAACAACA  1260

Qy  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT  1320

Qy  1321 ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATCCTTAAATCGTAGTTTTGGA  1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATCCTTAAATCGTAGTTTTGGA  1380

Qy  1381 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAAACTAATTGGGACAGATCCTAAC  1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 ATTGATAGTAAACAACTTCTGCAAACAAAATAGTGAAACTAATTGGGACAGATCCTAAC  1440

Qy  1441 GTTATTGAAGCAAGAATATATGCTCGTCGTGGTAATGATGAAGCTAACTAGGATTCAAATGAAATTCCAATA  1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTTATTGAAGCAAGAATATATGCTCGTCGTGGTAATGATGAAGCTAACTAGGATTCAAATGAAATTCCAATA  1500

Qy  1501 ACTAATGCAGGAAACTTTATCCGAAACACAACAATTGGTGTGTTTACTTCAACAGGT  1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTAATGCAGGAAACTTTATCCGAAACACAACAATTGGTGTGTTTACTTCAACAGGT  1560

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCAGCTCAACTGGATCAACGTCAACTGGAAACTTCCAA  1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TCAAGAGTAGTTTTAAGAGCTTCAGCTCAACTGGATCAACGTCAACTGGAAACTTCCAA  1620

Qy  1621 CCTTTCTTATACGTATTGGTTATTGGTTATTGGTATTGGATACCAACAACAGGAACTTTCTGG  1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CCTTTCTTATACGTATTGGTTATTGGTTATTGGTATTGGATACCAACAACAGGAACTTTCTGG  1680

Qy  1681 TACGGAACATATAAGCTTTTAAACAACAGCGCCCTTACGACGTATTGAGATTCTCCAAGAGT  1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACGGAACATATAAGCTTTTAAACAACAGCGCCCTTACGACGTATTAGATTCTCCAAGAGT  1740

Qy  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACTACTATATAAGAGAAGGTGGA  1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACTACTATATAAGAGAAGGTGGA  1800

Qy  1801 CTAACTGAAGAAGGTGCTAGAAGTCAAATATCCTTGCTATTCTAATAACTAATCGGAGAT 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CTAACTGAAGAAGGTGCTAGAAGTCAAATATCCTTGCTATTCTAATAACTAATCGGAGAT 1860

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT 1920

Qy  1921 CAATCAGTTTTAGGATTTGATGGAATTAGAAAATAACTTAAATGTTGGGGTTAAAGCATCA 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CAATCAGTTTTAGGATTTGATGGAATTAGAAAATAACTTAAATGTTGGGGTTAAAGCATCA 1980
```

```
Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAATCCGGTCTAGAAATGATTGCTGCAACACA 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAATCCAACGGTCTAGAAATGATTGCTGCAACACA 2040

Qy  2041 TACTTAAGATCACAAATGGATTAGCTAGAACATCTGATTACCAAACCAACAACCATTC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TACTTAAGATCACAAATGGATTAGCTAGAACATCTGATTACCAAACCAACAACCATTC 2100

Qy  2101 GGAACAACTCACCAAGTTATTTCAGTATCGATCGGTGATCAGTTCTCATCAATTAAGAAT 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 GGAACAACTCACCAAGTTATTTCAGTATCGATCGGTGATCAGTTCTCATCAATTAAGAAT 2160

Qy  2161 ATTAGAACAAATCTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATTAGAACAAATCTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220

Qy  2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGGTCAAGCTCA 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGGTCAAGCTCA 2280

Qy  2281 TTCAGTCCAACAAGTTTAATTGACGTTAATGAAAATTGTGTAATCTTACCTTTATTAGAC 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TTCAGTCCAACAAGTTTAATTGACGTTAATGAAAATTGTGTAATCTTACCTTTATTAGAC 2340

Qy  2341 AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTCTTCATCAAACCCTGGT 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 AATTCATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTCTTCATCAAACCCTGGT 2400

Qy  2401 TCTCCTGGATCATATACGCTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTT 2460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TCTCCTGGATCATATACGCTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTT 2460

Qy  2461 GAAGGTTCGTGGTGCTAAGTATACATCTGATTCTCTGGGGAACAATCCAATTCAAACCCGAT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GAAGGTTCGTGGTGCTAAGTATACATCTGATTCTCTGGGGAACAATCCAATTCAAACCCGAT 2520

Qy  2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAGAACTTCGTTACAAACCAA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAGAACTTCGTTACAAACCAA 2580

Qy  2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy  2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTACCTCTAAAAGTTCAGATCCAA 2700

Qy  2701 TACTTAGATGGTAAGTAGTATTATGATGATGCTAAATTAAAGAACAATAATTAGTAGCATTCTCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 TACTTAGATGGTAAGTAGTATTATGATGATGCTAAATTAAAGAACAATAATTAGTAGCATTCTCT 2760

Qy  2761 TATAACAACTTTGGCGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 TATAACAACTTTGGCGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Qy  2821 GGTATTCTTGCAATTATGATCATCTTAGGATATCGGTATTCCTTAAGAGCTTCAA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 GGTATTCTTGCAATTATGATCATCTTAGGATATCGGTATTCCTTAAGAGCTTCAA 2880

Qy  2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 AGAAAATTACAAGACAAGCAAAGGGTTCAAAACAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940

Qy  2941 GCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 GCTGTTGGTTCAGTTTACAAGAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAACCT 3000

Qy  3001 GCTGCTTTAGGTCGTGGTGCTAAATCGTGGTAAATCGGTAAAATCGGGTGATAAGAAACCTGCTGCT 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 GCTGCTTTAGGTCGTGGTGCTAAATCGTGGTAAATCGGTAAAATCGGGTGATAAGAAACCTGCTGCT 3060

Qy  3061 GCTCCAGCTAAACCATCTGCACCAAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 GCTCCAGCTAAACCATCTGCACCAAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
```

```
Db  3061 GCTCCAGCTAAACCATCTGCACCAAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAA 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCACCGCTCC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACCGCTCC 3180

Qy  3181 AAAGAATAA 3189
         |||||||||
Db  3181 AAAGAATAA 3189
```

RESULT 2
AAT75087
ID   AAT75087 standard; DNA; 3189 BP.
XX
AC   AAT75087;
XX
DT   27-FEB-1998 (first entry)
XX
DE   DNA encoding antigenic protein derived from Mycoplasma gallisepticum.
XX
KW   Mycoplasma gallisepticum antigen; epitope; monoclonal antibody; bird;
KW   mycoplasma infection; viral vaccine; avipoxvirus; herpesvirus; ss.
XX
OS   Mycoplasma gallisepticum.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..3189
FT   /*tag= a
FT   /product= "Mycoplasma_antigen"
FT   CDS   1..3189
FT   /*tag= b
FT   /codon= (seq:tga, aa:Trp)
XX
PN   WO9724370-A1.
XX
PD   10-JUL-1997.
XX
PF   27-DEC-1996;   96WO-JP003863.
XX
PR   28-DEC-1995;   95JP-00352754.
XX
PA   (JAPG) NIPPON ZEON KK.
XX
PI   Fujisawa A, Yoshida S;
XX
DR   WPI; 1997-363621/33.
DR   P-PSDB; AAW22721.
XX
PT   Antigenic protein derived from Mycoplasma gallisepticum - useful in
PT   vaccines against, and for diagnosis of mycoplasma infection in birds.
XX
PS   Claim 4; Page 24-31; 45pp; Japanese.
XX
CC   This sequence represents DNA encoding an antigenic protein of Mycoplasma
CC   gallisepticum having an epitope specifically recognised by the monoclonal
CC   antibody (MAb) 35A6. This antigenic protein inhibits the metabolism of
CC   bird infecting mycoplasma. The antigenic protein and the DNA molecule can
CC   be used in the preparation of component and viral vaccines against
CC   mycoplasma infection in birds. Mycoplasma infection can be diagnosed by
CC   hybridising DNA in a sample with a DNA probe corresponding to 10 or more
CC   (preferably 15 to 40) bases of the antigenic protein gene
XX
SQ   Sequence 3189 BP; 1130 A; 572 C; 522 G; 965 T; 0 U; 0 Other;

Query Match   98.6%; Score 3144.2; DB 2; Length 3189;
Best Local Similarity   99.1%; Pred. No. 0;
Matches 3161; Conservative   0; Mismatches   28; Indels   0; Gaps   0;

Qy   1 ATGAATATTTCTAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAATATTTCTAAAAAACTTAAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60

```
Qy    61  GCTCTTGGTTCTGCGAAGCTTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  GCTCTTGGTTCTGCGAAGCTTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120

Qy   121  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA  180

Qy   181  AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240

Qy   241  AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTAGATCTAAGT  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTAGATCTAAGT  300

Qy   301  GATGATTGTGGTGGATACAAGTAAAACAAAATAGTTTCAGATTACACAACTAGCAGAGAAAT  360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GATGATTGTTGCTGGATAACAAGTAAAACAAAATAGTTTCAGATTACACAACTAGCAGAGAAAT  360

Qy   361  AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGGTTAATGATGAAGCTAAC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGGTTAATGATGAAGCTAAC  420

Qy   421  GTTCATTTAAAAAGAATTAATACTAACTAACTCAAATAGAATTGGTAATAGAAACAACAATCT  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GTTCATTTAAAAAGAATTAATACTAACTAACTCAAATAGAATTGGTAATAGAAACAACAATCT  480

Qy   481  AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  AAGTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGG  540

Qy   541  ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATGTTAATGACTTCATTTTAGAT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  ACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATGTTAATGACTTCATTTTAGAT  600

Qy   601  GCGCCAATCTTACCCTAAAGATTTACACCCAGATTGGTATAACAACTTATACATTCAAAGAAAG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GCGCCAATCTTACCCTAAAGATTTACACCCAGATTGGTATAACAACTTATACATTCAAAGAAAC  660

Qy   661  ATCTTACCAAAATGACAGCTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTTGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  ATCTTACCAAAATGACAGCTCAACACTGCAGTTGTTCCTTGACCAGTAGGTAGAGTTAGTTGA  720

Qy   721  ACAAATGCTAATCAACCACTACTGATAATATGTTGGTCAAATAACTAATACAGATCCT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  ACAAATGCTAATCAACCACTACTGATAATATGTTGGTCAAATAACTAATACAGATCCT  780

Qy   781  ATTGCTCAAACTAAAACCACTACTGATAATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  ATTGCTCAAACTAAAACCACTACTGATAATAATCAAAATCCTTCAACTTTTAATTCAGGAGCA  840

Qy   841  ATGCCTGGTGCAAACAATAGATACGATTCAAATTGAATGTCAAGCATAGAATTAAAACA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  ATGCCTGGTGCAAACAATAGATACGATTCAAATTGAATGTCAAGCATAGAATTAAAACA  900

Qy   901  TCTTTCCAATTAGATGAAAAAATTGTTTATCCAGAAATGACTGGTTCTGAAGGAGAATAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TCTTTCCAATTAGATGAAAAAATTGTTTATCCAGAAATGAACTGGTTCTGAAGGAGAATAAA  960

Qy   961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  AATATTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGAATTCTTGAC  1020

Qy  1021  ATACCCGGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTATTTTCAAGACTA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  ATACCAGGGACTCCACAAGTTACTTAAAAGAAGAATTCAGTTAACGTATTTTCAAGACTA  1080

Qy  1081  TACTTAAACTCAGTTAATTCTTATCATTCATTGTTGATAGTATTTATATTTTTGGTACC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  TACTTAAACTCAGTTAATTCTTATCATTCATTGTTGATAGTATTTATATTTTTGGTACT  1140
```

```
Qy  1141  TCTGAATTACCATCATCATATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  TCTGAATTACCATCATCATATGATACTATTCATTCCCAACTAGATTATCTGATCTAACCGCT  1200

Qy  1201  TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA  1260

Qy  1261  AACGGAACAACGACAACAGCTGATACATCTGAATCTCAGTGGTCAACAGGTGCTGGAACAGGAAAT  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  AACGGAACAACGACAACAGCTGATACATCTGAATCTCAGTGGTCAACAGGTGCTGGAACAGGAAAT  1320

Qy  1321  ACTACTACACTTCTCAAACAGTTCTCAATCCTTAAAATCGTAGTTTGGA  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  ACTACTACACTTCTCAAACAGTTCTCAATCCTTAAAATCGTAGTTTTGGA  1380

Qy  1381  AATTGATAGTAAACCACTTCTGCAAACAAAATAGTGAAACTAAATTGGGCAGATCCTAAC  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  AATTGATAGTAAACCACTTCTGCAAACAAAATAGTGAAACTAAATTGAGCAGATCCTAAT  1440

Qy  1441  GTTATTGAAGCAAGAATATATGCTGAATACAGATAGGTATTCAAAATGAAATTCCAATA  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  GTTATTGAAGCAAGAATATATGCTGAATACAGATAGGTATTCAAAATGAAATTCCAATA  1500

Qy  1501  ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  ACTAATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGT  1560

Qy  1561  TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTGATCAACTCCAACTGGAAACTTCCAA  1620

Qy  1621  CCTTCTTATACGTATTTGGTATTTAGGATCAGACAGCCCTTAGGACGTATTAGATTTCTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  CCTTCTTATACGTATTTGGTATTTAGGATCAGACAGCCCTTAGGACGTATTAGATTTCTGA  1680

Qy  1681  TACGGAACATATAAGCTTTTAAACAACAGCCCTTAGGACGTATTAGATTCTCCAAGAGTA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TACGGAACATATAAGCTTTTAAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA  1740

Qy  1741  GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACATACACCGTGTTATGGTGGATAT  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  GGTACTGAAACCAATCAATTTAGAAGAACTTCATTAACATACACCGTGTTATGGTGGATAT  1800

Qy  1801  CTAACTGAAGAAGCTGCTAGAAGTTTCTCTAATACTCCATATAGAGACAACAAGGTGAC  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  CTAACTGAAGAAGCTGCTAGAAGTTTCTCTAATACTCCATATAGAGACAACAAGGTGAC  1860

Qy  1861  ACACCAGAAAGCCGAAGCAATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ACACCAGAAAGCCGAAGCAATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATT  1920

Qy  1921  CAATCAGTTTAGGATTTGATGGAATTAGAAATAACTTAAAGTTGGGGTTAAAGCATCA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CAATCAGTTTAGGATTTGATGGAATTAGAAATAACTTAAAGTTGGGGTTAAAGCATCA  1980

Qy  1981  AGCTTCTTAAACTCAAATAGAACCAAATCCAAACGGTCTAGAAATGATTCTGCAACAACA  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  AGCTTCTTAAACTCAAATAGAACCAAATCCAAACGGTCTAGAAATGATTCTGCAACAACA  2040

Qy  2041  TACTTAAGATCACAAAATTGGATTAGCTAGAAAACCAACACCATTC  2100
          |||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TACTTAAGATCACAAAATTGGATTAGCTAGAAAACCAACACCATTC  2100

Qy  2101  GGAACAACTCACCAAGTTATTTCAGTATCACCGGTATCAATTCATCAATTAGAAT  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  GGAACAACTCACCAAGTTATTTCAGTATCACCGGTATCAATTCATCAATTAGAAT  2160

Qy  2161  ATTAGAACAATCTTCCCTGGTAACCAGTATCGTATGGTACTTCTTATTCACAAATGAAAATAAT  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  ATTAGAACAATCTTCCCTGGTAACCAGTATGATATCCAGTTATGATACTTCTTATTCACAAATGAAAATAAT  2220

Qy  2221  AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2280
```

```
XX
DT   11-AUG-2003  (first entry)
XX
DE   DNA encoding Mycoplasma gallisepticum live vaccine.
XX
KW   Vaccine; ds; vaccination; cytadherence-deficiency; virucide.
XX
OS   Mycoplasma gallisepticum.
XX
PN   US2002187162-A1.
XX
PD   12-DEC-2002.
XX
PF   19-APR-2002; 2002US-00125818.
XX
PR   21-APR-2001; 2001US-0285569P.
XX
PA   (GEAR/) GEARY S J.
PA   (SILB/) SILBART L.
PA   (MARC/) MARCUS P.
PA   (SEKE/) SEKELLICK M.
XX
PI   Geary SJ, Silbart L, Marcus P, Sekellick M;
XX
DR   WPI; 2003-341017/32.
XX
PT   New method of avian vaccination against virulent strains of Mycoplasma
PT   gallisepticum by administering to a bird an immunogen comprising a
PT   cytadherence-deficient M. gallisepticum.
XX
PS   Claim 21; Page 6-10; 12pp; English.
XX
CC   This invention relates to a novel avian vaccination against virulent
CC   strains of Mycoplasma gallisepticum. The vaccination comprises
CC   administering to a bird an immunogen comprising a cytadherence-deficient
CC   M. gallisepticum having an inability to express at least two of three of
CC   the following proteins expressed by wild-type M. gallisepticum;
CC   cytadhesin molecule GapA; crmA protein; 45 kDa protein. The composition
CC   of the invention may have virucide activity and may be used as a
CC   virucide. The vaccine is useful for preventing M. gallisepticum virus
CC   infection in birds. The present sequence represents the DNA sequence
CC   encoding the Mycoplasma gallisepticum vaccine of the Gap-A, crmA and 45kDa
CC   sequence contains the coding sequences of the invention. This
CC   proteins which are not expressed in cytadherence deficient M.
CC   gallisepticum
XX
SQ   Sequence 8354 BP; 2916 A; 1463 C; 2545 G; 1426 T; 0 U; 4 Other;

Query Match             98.6%;  Score 3143.4;  DB 7;  Length 8354;
Best Local Similarity   99.1%;  Pred. No. 0;
Matches  3159;  Conservative    2;  Mismatches   28;  Indels    0;  Gaps    0;

Qy        1 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3778 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 3837

Qy       61 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3838 GCTCTTGGTTCTGCAAGCTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 3897

Qy      121 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCTGTTAGTACTTGCAGGTCTTGGACAA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3898 TTAGTTAATCAAGCAAGACGCTAGATGCTAATTCTGTTAGTACTTGCAGGTCTTGGACAA 3957

Qy      181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3958 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 4017

Qy      241 AATGGAACAATATCAAATTACAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4018 AATGGAACAATATCAAATTACAAATTAGATAGTTTACTAAACCATTATATGGTTAGATCTAAGT 4077

Qy      301 GATGATTGTGGTGGATACAAAGTAAAACAAAATAGTTTCAGATTACACAACTAGCAGAAAT  360
```

```
Db     2221 AAATCTAGTGTTTATACATTAGGTTAATGACGTTGACTCAGTAACCCTGATGCGTCAAGCTCA 2280
                                                       ||||||||||||
Qy     2281 TTCAGTCCAACAAGTTAATTGACGTTAATTGGTGTAATCTTTACCTTTTATTAGAC 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 TTCAGTCCAACAAGTTAATTGACGTTAATTGGTGTAATCTTTACCTTTTATTAGAC 2340

Qy     2341 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCAAACCCTGGT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 AATTCATTCTATACAGTAAATGCTGCTGGTAAATGTTGCATTGTTCTCATCAAACCCTGGT 2400

Qy     2401 TCTCCTGGATCATATACGTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 TCTCCTGGATCATATACGTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT 2460

Qy     2461 GAAGGTTCTGGTGCTGCTAAGTATACATCTCGATTTCTGGGGAACAATCCAATTCAAACCCGAT 2520
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 GAAGGTTCTGGTGCTGCTAAATATACATCTCGATTTCTGAGGAACAATCCAATTCAAACCCGAT 2520

Qy     2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAAACCAA 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAAACCAA 2580

Qy     2581 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTCGCTAAATGCTGGTACTAACTACCGTGTA 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTCGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy     2641 GTGGTTGATCCTGATGGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCATTA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 GTGGTTGATCCTGATGGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCATTA 2700

Qy     2701 TACTTAGATGGTAAGTAGTATTATGATGAGTCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 TACTTAGATGGTAAGTAGTATTATGATGAGTCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760

Qy     2761 TATAACAACTTGCGCGCCTACCTTCATGGGTAGTGCCTACAGCCAATTGGTAGTACATTA 2820
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TATAACAACTTGCGCGCTTTACCTTCATGAGTGTGCCTACAGCCAATTGGTAGTACATTA 2820

Qy     2821 GGTATTCTTGCAATTATGATGATCATCTTAGGATTAGCTATCGGTATCGGTATTCCTTTAGAGCTCAA 2880
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 GGTATTCTTGCAATTATGATGATCATCTTAGGATTAGCTATCGGTATCGGTATTCCTTTAGAGCTCAA 2880

Qy     2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940

Qy     2941 GCTGTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GCTGTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000

Qy     3001 GCTGCTTTAGGTGCGGTAAATCTGGTGATAAGAAGAAACCTGGTGCTGCTAAACCTGCT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 GCTGCTTTAGGTGCGGTAAATCTGGTGATAAGAAGAAACCTGGTGCTGCTAAACCTGCT 3060

Qy     3061 GCTCCAGCTAAACCATCTGCACCAAAAACCAACTGCTCCACCAGCTAAACCAACTGGGCCTAAA 3120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 GCTCCAGCTAAACCATCTGCACCAAAAACCAACTGCTCCACCAGCTAAACCAACTGGGCCTAAA 3120

Qy     3121 TCTGGTGCGCGCCTACAACCAACTGCTCCTAAGCCAGCTGCTCCTAAAACCAACCGCTCCC 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 TCTGGTGCGCGCCTACAACCAACTGCTCCTAAGCCAGCTGCTCCTAAAACCAACCGCTCCC 3180

Qy     3181 AAAGAATAA 3189
            ||| ||||
Db     3181 AAACAATAA 3189
```

```
RESULT 3
ABX95108
ID   ABX95108 standard; DNA; 8354 BP.
XX
AC   ABX95108;
```

```
Db  5158  ATTGATAGTAAACCAACTTCTGCAAACAAAATAGATGAAAATCAATTGAGCAGATCCTAAT  5217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1441  GTTATTGAAGCAAGAATATATGCTGAATACAGATTAGTAGTATTCAAAATGAAATTCCAATA  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5218  GTTATTGAAGCAAGAATATATGCTGAATACAGATTAGTAGTATTCAAAATGAAATTCCAATA  5277

Qy  1501  ACTAATGCAGGAAAACTTATCCGAAACACACAATTGGTGGTGTTGGTTTGGTTTACTTCAACAGGT  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5278  ACTAATGCAGGAAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTGGTTTACTTCAACAGGT  5337

Qy  1561  TCAAGAGTAGTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5338  TCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAA  5397

Qy  1621  CCTTTCTTATACGTATTGGTTATTTGGTTATTTAGGATACCAACAAAACTAGACAGGAACTTCTGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5398  CCTTTCTTATACGTATTGGTTATTTGGTTATTTAGGATACCAACAAAACTAGACAGGAACTTCTGA  5457

Qy  1681  TACGGACATATAAGCTTTAAACAACAGCCCCTTACGACGTATTAGATTCTCCAAGAGTA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5458  TACGGAACATATAAACTATTAAACAACAGCCCCTTACGACGTATTAGATGCTGCAAGAGTA  5517

Qy  1741  GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACATAACCCTGTTATGGGTGGATAT  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db  5518  GGTACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCAGTTATGGGTGGATAY  5577

Qy  1801  CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGGTGAC  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5578  CTAACTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGGTGAC  5637

Qy  1861  ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAAATACTTATGAGTACATT  1920
          |||||||||||||||||||||||||||||||||||||||||||||||| :
Db  5638  ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTAYTCTGATAAATACTTATGAGTACATT  5697

Qy  1921  CAATCAGTTTAGGATTTGATGAATTAGAAATAAACTTAAAATGTTGGGGTTAAAGCATCA  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5698  CAATCAGTTTAGGATTTGATGAATTAGAAATAAACTTAAAATGTTGGGGTTAAAGCATCA  5757

Qy  1981  AGCTTCTTAAACTCAAATAGACCAAATCAAACGGTCTGAGAAATGATTGCTGCAACAACA  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5758  AGCTTCTTAAACTCAAATAGACCAAATCAAACGGTCTGAGAAATGATTGCTGCAACAACA  5817

Qy  2041  TACTTAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAACCAACCAACCAACCATTC  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5818  TACTTAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAACCAACCAACCACCATTC  5877

Qy  2101  GGAACAACTCACCAAGTTATTTCAGTATCACCGGTGATCAGTTCTCATCATTAGGAAT  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5878  GGAACAACTCACCAAGTTATTTCAGTATCACCGGTGATCAGTTCTCATTCACCAAGGAAT  5937

Qy  2161  ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5938  ATTAGAACAATCTTCCCTGGTAACCAGTTATGATGATACTTCTTATTCACAAATGAAAATAAT  5997

Qy  2221  AAATCTAGTGTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5998  AAATCTAGTGTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA  6057

Qy  2281  TTCAGTCCAACAGTTTAATTGACGTTAATGAAATTGGTAATCTTTATTAGAC  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6058  TTCAGTCCAACAGTTTAATTGACGTTAATGAAATTGGTAATCTTTATTAGAC  6117

Qy  2341  AATTCATTCTATACAGTAAATGCTGCTGTAAATACATTTAATCAGAACTTAATCAGATATTGCTTTT  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6118  AATTCATTCTATACAGTAAATGCTGCTGTAAATACATTTAATCAGAACTTAATTGCTTTT  6177

Qy  2401  TCTTCCTGGATCATATACTGCTGTAAATACTTATCGTAGTTTTGGA  2460
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  6178  TCTTCCTGGATCATATACTGCTGTAAATACTTATCGTAGTTTTGGA  6237

Qy  2461  GAAGGTTCGGTGCTAAGTACATCTGATTTCTGGGAACAATCCAATTCAAACCCGAT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  6238  GAAGGTTCGGTGCTAAATATACATCTGATTTCTGAGGAACAATCCAATTCAAACCCGAT  6297
```

```
Db  4078  GATGATTTTGCTGGATACAAACAAATAGTTCAGATTACACAACTAGCAGAAAT  4137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  361   AGATTTGATCAAGACAAAACAAGAGAGCATATCTCTGTTGGTTAATGAAGCTAAC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4138  AGATTTGATCAAGACAAAACAAGAGCATATCTCTGTTGGTTAATGAAGCTAAC  4197

Qy  421   GTTCATTTAAAAAGAATTAATACTAACTAATGGTAATAGAATAGAAACAACAATTCT  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4198  GTTCATTTAAAAAGAATTAATACTAACTAATGGTAATAGAATAGAAACAACAATTCT  4257

Qy  481   AAGTTTGTAATTGGTGGTGTTGGTGTTGATAATCAGCTCACGTAATTACTGATGATGGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4258  AAGTTTGTAATTGGTGGTGTTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGG  4317

Qy  541   ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAATGTTAATGACTTCATTTTAGAT  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4318  ACTAAATTTAATTTTACAAACAAAGCAAACTCAAGGTGAAATGTTAATGACTTCATTTTTAGAT  4377

Qy  601   GCGCCAATCTTACCTAAAGATTACACCCGAGATTGGTATAACTAACTTATACATTCAAAGAAAG  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4378  GCGCCAATCTTTACCTAAAGATTACACCCGAGATTGGTATAACTAACTTATACATTCAAAGAAAG  4437

Qy  661   ATCTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGA  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4438  ATCTTTACCAAATGACGTAAACACTGCAGTTGTTCCTTGACCAGTAGGTAGAGTTAGTGGA  4497

Qy  721   ACAAATGCTGATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4498  ACAAATGCTGATGATGGGATGTTGATTGTTTGGGAATGGTCAAATAACTAATAACAGATCCT  4557

Qy  781   ATTGCTCAAACTAAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCA  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4558  ATTGCTCAAACTAAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCA  4617

Qy  841   ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAAATGTCAAGCATAGAAATTAAAACA  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4618  ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAAATGTCAAGCATAGAAATTAAAAACA  4677

Qy  901   TCTTTCCAATTGAATGAAAAATTTGTTTATCCAGAATCTGGTTCTGAAGAGAATAAA  960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4678  TCTTTCCAATTGAATGAAAAATTTGTTTATCCAGAATCTGGTTCTGAAGAGAATAAA  4737

Qy  961   AATATTACAAGATTACAGATGATTAGCTACTGGAAGTTGCCAAGCAACAACGAAAGATATTGGATTCTTGAC  1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4738  AATATTACAAGATTACAGATGATTAGCTACTGGAAGTTGCCAAGCAACAACGAAAGATATTGGAATTCTTGAC  4797

Qy  1021  ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTTCAAGACTA  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4798  ATACCAGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTAATTTTCAAGACTA  4857

Qy  1081  TACTTAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTTATATTTTTGGTACC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4858  TACTTAAACTCAGTTAATTCTTTATCATTGGTGATAGTATTTATATTTTTGGTACT  4917

Qy  1141  TCTGAATTACCATCATTATGTGTACTATTATTGATTATCTGATCTAACCGCT  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4918  TCTGAATTACCATCATTATGTGTACTATTATTGATTATCTGATCTAACCGCT  4977

Qy  1201  TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACCGTACAACAACA  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4978  TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACCGTACAACAACA  5037

Qy  1261  AACGGAACAACGACAACAGAGCTGATCATCTCAGTGGTTCAACAGTGCTGGAACAGGAAAT  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5038  AACGGAACAACGACAACAGAGCTGATCATCTCAGTGGTTCAACAGTGCTGGAACAGGAAAT  5097

Qy  1321  ACTACTAACACTTCTCAAAACAGTTTCTAATCCTACTTAAATATACTTATCGTAGTTTTTGGA  1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5098  ACTACTAACACTTCTCAAAACAGTTTCTAATCCTACTTAAATATACTTATCGTAGTTTTTGGA  5157

Qy  1381  ATTGATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAAC  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Qy 2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6298 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 6357

Qy 2581 AGCTTCCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6358 AGCTTCCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 6417

Qy 2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6418 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAA 6477

Qy 2701 TACTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6478 TACTTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAAATTTAGTAACATTCTCT 6537

Qy 2761 TATAACAACTTTGGCGCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6538 TATAACAACTTTCGTGCTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTAGTACATTA 6597

Qy 2821 GGTATTCTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAGAGCTCAA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6598 GGTATTCTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAGAGCTCAA 6657

Qy 2881 AGAAAATTACAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6658 AGAAAATTACAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 6717

Qy 2941 GCTGTTGGTTCAGTTACAAGAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6718 GCTGTTGGTTCAGTTACAAGAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAACCT 6777

Qy 3001 GCTGCTTTAGGTGCTGGTAAATCGGTAAGAAGAAACCTGCTGCTGCTAAACCTGCT 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6778 GCTGCTTTAGGTGCTGGTAAATCGGTAATCGGTAAGAAGAAACCTGCTGCTGCTAAACCTGCT 6837

Qy 3061 GCTCCAGCTAAACCATTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6838 GCTCCAGCTAAACCATTCGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGCGCCTAAA 6897

Qy 3121 TCTGGTGCGCCTACAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCACCGCTCCC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 6898 TCTGGTGCGCCCTACAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCACCGCTCCC 6957

Qy 3181 AAAGAATAA 3189
         |||||||||
Db 6958 AAAGAATAA 6966

RESULT 4
AAF88025
ID AAF88025 standard; DNA; 10651 BP.
XX
AC AAF88025;
XX
DT 16-JUL-2001 (first entry)
XX
DE Vector pcDNA3-FasL-IRES-crmA DNA.
XX
KW Vector; pcDNA3-FasL-IRES-crmA; autoimmune disease; apoptosis; crmA;
KW suicide enzyme; antiinflammatory; immunomodulatory; rheumatic; arthritic;
KW dermatological; gene therapy; rheumatoid arthritis;
KW systemic lupus erythematosus; Sjoegren-Syndrome; polymyositis;
KW dermatomyositis; polymyalgica; rheumatism; arteritis temporalis;
KW Crohn's disease; Bechterew disease; colitis ulcerosa; thyroiditis;
KW autoimmune hepatitis; diabetes mellitus (type I); psoriasis; dermatitis;
KW pemphigus vulgaris; multiple sclerosis; myasthenia gravis;
KW viral infection; brain inflammation; transplant rejection; ds.
XX
OS Synthetic.
XX
PN WO200127254-A2.

XX
PD 19-APR-2001.
XX
PF 12-OCT-2000; 2000WO-DE003608.
XX
PR 12-OCT-1999; 99DE-01048983.
XX
PA (SCHW/) SCHWARZMANN F.
XX
PI Schwarzmann F;
XX
DR WPI; 2001-282021/29.
XX
PT New gene transfer vector useful in the prevention and treatment of
PT autoimmune and chronic inflammatory disease.
XX
PS Claim 9; Page 74-79; 82pp; German.
XX
CC This invention describes a novel gene transfer vector (N1) which
CC comprises (1) a first nucleic acid sequence (I) encoding one or more
CC ligands that trigger apoptosis; (2) a second nucleic acid sequence (II)
CC encoding one or more antigens; and optionally (3) a third nucleic acid
CC sequence (III) encoding one or more anti-apoptosis molecules; and
CC optionally (4) a fourth nucleic acid sequence (IV) encoding one or more
CC suicide enzymes. The products of the invention have antiinflammatory,
CC immunomodulatory, rheumatic, arthritic and dermatological activity and
CC can be used in gene therapy. The gene vector may be used to produce
CC medicine for the prevention or therapy of autoimmune diseases, such as
CC rheumatoid arthritis, systemic lupus erythematosus, Sjoegren-Syndrome,
CC polymyositis, dermatomyositis, polymyalgica, rheumatism, arteritis
CC temporalis, Crohn's disease, Bechterew disease, colitis ulcerosa,
CC autoimmune hepatitis, diabetes mellitus (type I), suprarena body
CC insufficiency, thyroiditis, psoriasis, dermatitis, pemphigus vulgaris,
CC multiple sclerosis, myasthenia gravis, or for the prevention or therapy
CC of diseases associated with chronic inflammation, that target
CC immunopathogens, particulary inflammations associated with viral or
CC bacterial infection, particularly hepatitis B or C infections or of the
CC inflammation of the brain after infection from Masern virus, or
CC transplant rejections. The vector may be used in ex vivo modification of
CC animal or mammal cells, particularly human cells. This sequence
CC represents the vector pcDNA3-FasL-IRES-crmA which contains the suicide
CC enzyme crmA
XX
SQ Sequence 10651 BP; 2782 A; 2618 C; 2490 G; 2759 T; 0 U; 2 Other;

Query Match   98.6%; Score 3143.4; DB 5; Length 10651;
Best Local Similarity   99.1%; Pred. No. 0;
Matches 3159; Conservative   2; Mismatches 28; Indels   0; Gaps   0;

Qy   1 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6 ATGAATATTTCTAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 65

Qy   61 GCTCCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 GCTCCTTGGTTCTGCAAGCTTTGGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAA 125

Qy 121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCAGGTCTTGGACAA 185

Qy 181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 186 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT 245

Qy 241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 246 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT 305

Qy 301 GATGATTGGTGGTGGATACAAAGTAAAACAAAAGTTTCAGATTACACAACTAGCAGAAAT 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 306 GATGATTTGTGGTGGATACAAAGTAAAACAAAAGTTTCAGATTACACAACTAGCAGAAAT 365

```
Qy   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCGTCGTTGGTTAATGATGAAGCTAAC  420
         ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   366 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTCGTTGGTTAATGATGAAGCTAAC  425

Qy   421 GTTCATTAAAAGAGAATTAATACTAACTAACTAATATAGAAATTGGTAATAGAAACAACAATTCT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   426 GTTCATTAAAAGAGAATTAATACTAACTAACTAATATAGAAATTGGTAATAGAAACAACAATTCT  485

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTACTGATGATGGG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   486 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTACTGATGATGGG  545

Qy   541 ACTAAATTTAATTTACAAACCAAACTCAAGGTGAAAATGTTAATGACTTCATTTAGAT  600
         |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   546 ACTAAATTTAATTTACAAAGCAAACTCAAGGTGAAAATGTTAATGACTTCATTTAGAT  605

Qy   601 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   606 GCGCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  665

Qy   661 ATCTTACCAAATGAGGTCAACACTGCAGTTGTTGATTGTGGGAATGGTCAGTAGAGTTAGTGA  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   666 ATCTTACCAAATGAGGTCAACACTGCAGTTGTTGATTGTGGGAATGGTCAGTAGAGTTAGTGA  725

Qy   721 ACAAATGCTGATGATGGGATGGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   726 ACAAATGCTGATGATGGGATGGTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCT  785

Qy   781 ATTGCTCAAACTAAAACCACTACTACTGATAATCCTCAACTTTTAATTCAAGCATAATGTCAAGCATAGGAGCA  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   786 ATTGCTCAAACTAAAACCACTACTACTGATAATCCTCAACTTTTAATTCAAGCATAATGTCAAGCATAGGAGCA  845

Qy   841 ATGCCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGTCAAGCAATTAAAACA  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   846 ATGCCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGTCAAGCAATTAAAACA  905

Qy   901 TCTTTCCAATTAGATGAAAAATTTGTTTTATCCAGAATGGACTCGGTTCTGAAGAGAATAAA  960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   906 TCTTTCCAATTAGATGAAAAATTTGTTTTATCCAGAATGGACTCGGTTCTGAAGAGAATAAA  965

Qy   961 AATAATTACAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   966 AATATTACAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGAATTCTTGAC  1025

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1026 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTA  1085

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTAGTATTTATATTTTGGTACC  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1086 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTAGTATTTATATTTTGGTACT  1145

Qy  1141 TCTGAATTCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGCTGATAATCTGATCTAACCGCT  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1146 TCTGAATTCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGCTGATAATCTGATCTAACCGCT  1205

Qy  1201 TTGAATCAAGTTAAAACAACGACAACAACTGATAGTGGTTCAACAGGTGCTGGAAACAGGAAAT  1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1206 TTGAATCAAGTTAAAACAACGACAACAACTGATAGTGGTTCAACAGGTGCTGGAAACAGGAAAT  1265

Qy  1261 AACGGAACAACGACAACGACAACTGATAGTATTGAAGCTTCAAGCACTGATAAACGTGAAACTAACTAAC  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1266 AACGGAACAACGACAACGACAACTGATAGTATTGAAGCTTCAAGCACTGATAAACGTGAAACTAACTAAC  1325

Qy  1321 ATACTAACACTTCTCAAACAGTTTCTCAAACAGTTTCTAAATCCTACTTAAATACTTATCGTAGTTTTGGA  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1326 ATACTAACACTTCTCAAACAGTTTCTCAAACAGTTTCTAAATCCTACTTAAATACTTATCGTAGTTTTTGGA  1385

Qy  1381 ATTGATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAAC  1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1386 ATTGATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGAGCAGATCCTAAT  1445

Qy  1441 GTTATTGAAGCAAGAATAATATGTCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||


Db  1446 GTTATTGAAGCAAGAATAATATGTCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA  1505

Qy  1501 ACTAATGCGAGGAAACTTTATCCGAAACAACACAATTGGTGCGTGTTGGTTTACTTCAACAGGT  1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1506 ACTAATGCGAGGAAACTTTATCCGAAACAACACAATTGGTGCGTGTTGGTTTACTTCAACAGGT  1565

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCTTCTTATAAACGTGATCAACGTCCAACTGGAAACTTCCAA  1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1566 TCAAGAGTAGTTTTAAGAGCTTCTTCTTATAAACGTGATCAACGTCCAACTGGAAACTTCCAA  1625

Qy  1621 CCTTTCTTATACGTATTGGTTATTGGTTATTAGGATACCAACCAGAACAGGAGAACTTTCTGG  1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1626 CCTTTCTTATACGTATTTGGTTATTGGTTATTAGGATACCAACCAACAAACTAGAGAACTTTCTGA  1685

Qy  1681 TACGGAACATATAAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA  1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1686 TACGGAACATATAAACTTTAAACAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTA  1745

Qy  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACATACACCCTGTTATGGGTGGATAT  1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db  1746 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACATACACCCTGTTATGGGTGGATAY  1805

Qy  1801 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCGGCTATTCTGATAATACTTATGAGTACATT  1860
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  1806 CTAACTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCGGCTAYTCGATAYTCGATAATACTTATGAGTACATT  1865

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATAACTTAAATGTTGGGGTTAAAGCATCA  1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1866 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTAYTCGATAYTCGATAATAACTTAAATGTTGGGGTTAAAGCATCA  1925

Qy  1921 CAATCAGTTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA  1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1926 CAATCAGTTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCA  1985

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1986 AGCTTCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACA  2045

Qy  2041 TACTTAAGATCACAAAATTGGATTAGCTGAATCTCGGATTACCAACCAACAACCATTC  2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2046 TACTTAAGATCACAAAATTGGATTAGCTGAATCTCGGATTACCAACCAACAACCATTC  2105

Qy  2101 GGAACAACTCACCAGTTATTTCGATGATCAGTCACCTGGTGATCAGTCTCTCATCAATTAAGAAT  2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2106 GGAACAACTCACCAGTTATTTCGATGATCACCTGGTGATCAGTCTCTCATCAATTAAGAAT  2165

Qy  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT  2220
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db  2166 ATTAGAACAATCTTCCCTGGTAACCAGTTATGTAGATCTTCTTATTCACAAATGAAAATAT  2225

Qy  2221 AAATCTAGTGTTTATACATTAAGATTGACTCAAGTAACCCTGATGGCGTCAAGCTCA  2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2226 AAATCTAGTGTTTATACATTAAGATTGACTCAAGTAACCCTGATGCGTCAAGCTCA  2285

Qy  2281 TTCAGTCCAACAGTTAATTGAGACGTTAATGAATTGGTGTAATCTTTATTAGAC  2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2286 TTCAGTCCAACAAGTTAATTGAGACGTTAATGAATTGGTGTAATCTTTATTAGAC  2345

Qy  2341 AATTCATTCTATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT  2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2346 AATTCATTCTATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTT  2405

Qy  2401 TCTCCTGGATCATATACTTGCTGTGTAAATCAGAGTTTTGGA  2460
         ||||||||||||||||||||||||||||||||||||||||||
Db  2406 TCTCCTGGATCATATACTTGCTGTGTAAATCAGAGTTTTGGA  2465

Qy  2461 GAAGGTTCTGGTGTCGTAAGTATACAACAATCCAATTCAAACCCGAT  2520
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2466 GAAGGTTCTGGTGTCGTAAATACATCGTAAATTCAGAGGAACAATTCCAATTCAAACCCGAT  2525

Qy  2521 GAGTACTTAATTCAAAATGGTTCACTAGTCAGTCAGTGGCTAGAAACTTCGTTACAAACCAA  2580
```

```
PI  Okuda T, Saito S, Dorsey KM, Tsuzaki Y;
XX
DR  WPI; 2003-373746/36.
DR  P-PSDB; ABR57376.
XX
PT  DNA molecule derived from a prokaryotic cell, useful for producing a
PT  vaccine for treating viral infections comprises at least one modified DNA
PT  regions encoding NXB so that no N-glycosylation occurs during expression.
XX
PS  Example 2; Page 62-64; 70pp; English.
XX
CC  The present invention describes a DNA molecule derived from a prokaryotic
CC  cell, where at least one of the DNA regions encoding NXB (where N =
CC  asparagine, X = any amino acid other than proline, and B = serine or
CC  threonine) has been modified so that no N-glycosylation occurs during the
CC  expression in a eukaryotic cell. Also described: (1) a fused DNA
CC  molecule, where a DNA encoding a signal sequence has been ligated to the
CC  N-terminal end of the modified DNA molecule as described above so that it
CC  may be expressed as a fusion protein; (2) a recombinant virus integrated
CC  with the DNA molecule or the fused DNA molecule described above; (3)
CC  producing a modified or fusion protein by using the recombinant virus
CC  described above, to express a protein encoded by the modified DNA
CC  molecule or the fused DNA molecule in a eukaryotic cell; and (4) a
CC  vaccine comprising the recombinant virus. The DNA molecule has virucide
CC  and immunostimulant activities. The DNA molecule is useful for producing
CC  a vaccine for treating viral infections. The present sequence is used in
CC  the exemplification of the present invention
XX
SQ  Sequence 3129 BP; 1078 A; 581 C; 541 G; 929 T; 0 U; 0 Other;

Query Match     95.9%;  Score 3057.8;  DB 7;  Length 3129;
Best Local Similarity   98.7%;
Matches 3083;   Conservative   0;  Mismatches   42;   Indels   0;   Gaps   0;

Qy    65 TTGGTTCGCAAGCTTTGGCTTTAAGCAATCAGCATCAGATAGAGTAACGATAACACGCAATTAG 124
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5 TCGGTTCGCAAGCTTTGGCTTTAAGCAATCAGCATCAGATAGAGTAACGATAACACGCAATTAG 64

Qy   125 TTAATCAAGCAAGAACGCTAGAGCTAATTCTGTTGCAGTGTCTTGGTTAATGATGAGGCTAACGTTC 184
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    65 TTAATCAAGCAAGAACGCTAGAGCTAATTCTGTTGCAGTGTCTTGGTTAATGATGAGGCTAACGTTC 124

Qy   185 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCTAATG 244
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   125 GTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTATAACAGCAGCTAATG 184

Qy   245 GAACAATTATCAAATTAGATAGTTTTACTAAAGCATTATATGGTTTAGATCTTAAGTGATG 304
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   185 GAACAATTATCAAATTAGATAGTTTTACTAAAGCATTATATGGTTTAGATCTTAAGTGATG 244

Qy   305 ATTGTGGTGGATCAAAGTGGATAAAACAAATAGTTCAGATTACAACTAGCAGAAATAGAT 364
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   245 ATTTTGCTGGATCAAAGTGGATAAAACAAATAGTTCAGATTACAACTAGCAGAAATAGAT 304

Qy   365 TTGATCAAGACAAGACAAGAGCATATATGCTCGTGTTGGTTAATGATGAAGGCTAACGTTC 424
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   305 TTGATCAAGACAAGACAAGAGCATATATGCTCGTGTTGGTTAATGATGAAGGCTAACGTTC 364

Qy   425 ATTTAAAAAGATTAATTAACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGT 484
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   365 ATTTAAAAAGATTAATTAACTAACTCAAATAGAATTGGTAATAGAACCAAAATTCTAAGT 424

Qy   485 TTGTAATTGGTGGTGTTGTTGATAATCCAGCTCACGTAGTTACTGATGATGGGACTA 544
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   425 TTGTAATTGGTGGTGTTGTTGATAATCCAGCTCACGTAGTTACTGATGATGGGACTA 484

Qy   545 AATTTAATTTACAAACCAAACTCAAGTGAAATTGTTAATGACTTCATTTTAGATGCGC 604
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   485 AATTTCAGTTACACAGCAAACTCAAGTGAAATTGTTAATGACTTCATTTTAGATGCGC 544

Qy   605 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATCATTCAAAGAAAGATCT 664
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   545 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATCATTCAAAGAAAGATCT 604
```

```
Db  2526 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2585
Qy  2581 AGCTTCTTAAAACAGTTTAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTACCGTGTA 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2586 AGCTTCTTAAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTACCGTGTA 2645

Qy  2641 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2646 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAA 2705

Qy  2701 TACTTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTTAACATTCTCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2706 TACTTAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTTAGTTAACATTCTCT 2765

Qy  2761 TATAACAACTTGCGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2766 TATAACAACTTGCTGCTTTACCTTCATGAGTAGTGCCTACAGCAATTGGTAGTACATTA 2825

Qy  2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCCAA 2880
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2826 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCCAA 2885

Qy  2881 AGAAAATTACAAGACAAGAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2886 AGAAAATTACAAGACAAGAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCT 2945

Qy  2941 GCTGTTGGTTCAGTTTACAAGAAGAATTATTACCCCAAACTGCTAACGTTAAGAAAAAACCT 3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2946 GCTGTTGGTTCAGTTTACAAGAAGAATTATTACCCCAAACTGCTAACGTTAAGAAAAAACCT 3005

Qy  3001 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTCGTGCTCGTAAACCTGCT 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3006 GCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTCGTGCTCGTAAACCTGCT 3065

Qy  3061 GCTCCAGCTAAACCATCTGCCACCAAAAGCTAGCTCACCGAGCTCAAACCAACTGGGCCTAAA 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3066 GCTCCAGCTAAACCATCTGCCACCAAAAGCTAGCTCACCGAGCTCAAACCAACTGCGCCTAAA 3125

Qy  3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3126 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3185

Qy  3181 AAAGAATAA 3189
         |||||||||
Db  3186 AAAGAATAA 3194
```

```
RESULT 5
ACF03435
ID   ACF03435 standard; DNA; 3129 BP.
XX
AC   ACF03435;
XX
DT   11-SEP-2003   (first entry)
XX
DE   Mycoplasma gallisepticum modified mgc3 gene SEQ ID NO:79.
XX
KW   DNA molecule; prokaryotic cell; eukaryotic cell; virucide; vaccine;
KW   immunostimulant; viral infection; gene; ds.
XX
OS   Mycoplasma gallisepticum.
XX
PN   EP1275716-A2.
XX
PD   15-JAN-2003.
XX
PF   11-JUL-2002; 2002EP-00254879.
XX
PR   11-JUL-2001; 2001US-00901572.
PR   25-APR-2002; 2002US-00131591.
XX
PA   (JAPG ) ZEON CORP.
XX
```

Qy 665 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGGAACAA 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 605 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGGAACAA 664

Qy 725 ATGCTGATGATGGGATGTTTGATTGTGTGGGAATGCTCAAATAACTAATACAGATCCTATTG 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 665 ATGCTGATGATGGGATGTTTGATTGTGTGGGAATGGTCAAATAACTAATACAGATCCTATTG 724

Qy 785 CTCAAACTAAAACCACTACTGATAATCAAAATCCTCAACTTTTAATTCAGGAGCAATGC 844
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 725 CTCAAACTAAAACCACTACTGATAATCAAAATCCTCAACTTTTAATTCAGGAGCAATGC 784

Qy 845 CTGGTGCAACAATAGATACGATCCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 904
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 785 CTGGTGCAACAATAGATACGATCCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 844

Qy 905 TCCAATTAGATGAAAAATTGTTATCCAGAATGACTGGTTCTGAAGAATAAAAAATA 964
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 845 TCCAATTAGATGAAAAATTGTTATCCAGAATGACTGGTTCTGAAGAATAAAACAGA 904

Qy 965 TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 1024
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 905 TTACAAGATTAGCTACTGGAAGTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 964

Qy 1025 CCGGGACTCCACAAGTTACTTTAAAAGAAGATCAGTTAACGTATTTTCAAGACTATACT 1084
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 965 CCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTATACT 1024

Qy 1085 TAAACTCAGTTAATTCTTTATCATTCATTCCCAACTAGATTATCGATCTAACCGCTTCTG 1144
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1025 TAAACTCAGTTAATTCTTTATCATTCATTCCCAACTAGATTATCGATCTAACCGCTTCTG 1084

Qy 1145 AATTACCATCATTATGTGACTATTCATTCAGATGAAGCTTAAAACGATGATATTGAAGC 1204
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1085 AATTACCATCATTATGTGACTATTCATTCAGATGAAGCTTAAAACGATGATATTGAAGC 1144

Qy 1205 ATCAGTTAAACAGATGATATTGAAGCTTCAAGCCACTGCATGAAGCTTCAAGCAACAAACG 1264
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1145 ATCAGTTAAACAGATGATATTGAAGCTTCAAGCCACTGCATGAAGCTTCAAGCAACAACAGG 1204

Qy 1265 GAACAACGACAACAGCTGATACATCTGAGTTGTTCAACAGGTGCTGGAACAGGAAATACTA 1324
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1205 GAACAACGACAACAGCTGATACATCTGAGTTGTTCAACAGGTGCTGGAACAGGACAGACTA 1264

Qy 1325 CTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1384
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1265 CTCGACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1324

Qy 1385 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGCAGATCCTAACGTTA 1444
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1325 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGCAGATCCTAACGTTA 1384

Qy 1445 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1504
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1385 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1444

Qy 1505 ATGCCAGGAAACTTATCCGAAACACAATTGGTGTGTGTTTACTTCAACAGGTTCAA 1564
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1445 ATGCCAGGAAACTTATCCGAAACACAATTGGTGTGTGTTTACTTCAACAGGTTCAA 1504

Qy 1565 GAGTAGTTTTAAGAGCTTCTTATAACGGGTATCAACGTGATCATCCTGAAACTTCAACCTT 1624
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1505 GAGTAGTTTTAAGAGCTTCTTATAACGGGTATCAACGTGATCATCCTGAAACTTCAACCTT 1564

Qy 1625 TCTTTATACGTATTTGTTATTTAGGATACCAACAAGATGAACAAAACTAGAACACAGGAACTTTCTTGGTACG 1684
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1565 TCTTTATACGTATTTGTTATTTAGGATACCAACAAGATGAACAAAACTAGAACACAGGAACTTTCTTGGTACG 1624

Qy 1685 GAACATATAAGCTTTTAAACAACAGGCCCTTACGACGTATTAGATGCTGCAAGAGTAGGTA 1744
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1625 GAACATATAAGCTTTTAACAAC-ACGAACAGCCCTTACGACGTATTAGATGCTGCAAGAGTAGGTA 1684

Qy 1745 CTGAAACCAATCAATTTAGAGAACTTCATTAACACATACCCTGTTATGGGTGATATCTAA 1804
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1685 CTGAAACCAATCAATTTAGAGAACTTCATTAACACATACCCTGTTATGGGTGATATCTAA 1744

Qy 1805 CTGAAGAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGTGACACAC 1864
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1745 CTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGTGACACAC 1804

Qy 1865 CAGAAGCCGAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1924
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1805 CAGAAAGCCGAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1864

Qy 1925 CAGTTTTAGGATTTGATGGAATTAGGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1984
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1865 CAGTTTTAGGATTTGATGGAATTAGGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1924

Qy 1985 TCTTAAACTCAAATGACCAAATCCAAAATGAAAATCGTTCTAGAAATGATTGCTGCAACAACATACT 2044
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1925 TCTTAAACTCAAATGACCAAATCCAAAATGAAAATCGTTCTAGAAATGATTGCTGCAACAACATACT 1984

Qy 2045 TAAGACAACAAATTGGATTAGCTAGCAACATCTGGATTACCAAACCAACAACCATTCGGAA 2104
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1985 TAAGACAACAAATTGGATTAGCTAGCAACATCTGGATTACCAAACCAACAACCATTCGGAA 2044

Qy 2105 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCCAGTTCTCATCAATTAAGAGAATATTA 2164
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2045 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCCAGTTCTCATCAATTAAGAGAATATTA 2104

Qy 2165 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAATAAAT 2224
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATCAGAAAT 2164

Qy 2225 CTAGTGTTTATACATTAAGATTAAGCTTAATGAAGTTAATGAAATTGGTGTAATCTTACTTTATTGAGACAATT 2284
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2165 CTAGTGTTTATACATTAAGATTAAGCTTAATGAAGTTAATGAAATTGGTGTAATCTTACTTTATTGAGACAATT 2224

Qy 2285 GTCCAACAAGTTTAATTGACGTTAATGAAGTTAATGAAATTGGTGTAATCTTACTTTATTGAGACAATT 2344
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2225 GTCCAACAAGTTTAATTGACGTTAATGAAGTTAATGAAATTGGTGTAATCTTACTTTATTGAGACAATT 2284

Qy 2345 CATTCTATACAGTAAATGCTGCTGGTAAATACATTGACTGGTAATGTTGCATTGTTGCATCATCCTGGTTCTC 2404
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2285 CATTCTATACAGTAAATGCTGCTGGTAAATACATTGACTGGTAATGTTGCATTGTTGCATCATCCTGGTTCTC 2344

Qy 2405 CTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTGAAG 2464
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2345 CTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTGAAG 2404

Qy 2465 GTTCTGGTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGATGAGT 2524
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2405 GTTCTGGTGCTAAGTATACATCTGATTTCTGGGGACAATCCAATTCAAACCCGATGAGT 2464

Qy 2525 ACTTAATTCAAAATGGGTTCACTGGTCAAGTGCCTAGAAACTTCGTTACAAACCAAAGCT 2584
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2465 ACTTAATTCAAAATGGGTTCACTGGTCAAGTGCCTAGAAACTTCGTTACACAGCCAAAGCT 2524

Qy 2585 TCTTAAACAGTTGAGTTGACTTCACTCCCGTACTAACCTACCGTGTAGTGG 2644
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2525 TCTTAAACAGTTGAGTTGACTTCACTCCCGTACTAACCTACCGTGTAGTGG 2584

Qy 2645 TTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAATACT 2704
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2585 TTGATCCTGATGGTAATTTAACAAACCAAAACCTACCTTCTAAAAGTTCAGATCCAATACT 2644

Qy 2705 TAGATGGTAAGTAGATTATGATGCTGAAATTAAAGAACAATAATTTAGTAACATTCTTATA 2764
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2645 TAGATGGTAAGTAGATTATGATGCTGAAATTAAAGAACAATAATTTAGTAACATTCTTATA 2704

Qy 2765 ACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA 2824
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2705 ACAACTTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA 2764

Qy 2825 TTCTTGCAATTATGATCATCTTAGGATTAGATTAGCTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAA 2884

CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product, or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 3159 BP; 1017 A; 639 C; 647 G; 856 T; 0 U; 0 Other;

Query Match   3.9%;   Score 124.6; DB 7;   Length 3159;
Best Local Similarity   63.5%;
Matches 190; Conservative   0; Mismatches 109;   Indels   0; Gaps   0;

Qy   2770 TTTGGGGCCTACCTTCATGGGTAGTGCCTACACGCAATTGGTAGTACATTAGGTATTCTT 2829
           ||||| |||| |||||| |||| | || | |||| |||| || || ||||| |||| |
Db   2797 TTTGCAGCACTCCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG 2856

Qy   2830 GCAATTATGATCATCTTAGGATTAGCTACGGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
           || | | ||| | ||| |||| |||| ||||| ||| || || || |||| ||
Db   2857 TTTATCTTGTTAGTCTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 2916

Qy   2890 CAAGACAAAGGGTTCAAAACAACATTCAAAAAGTGATACCTTGACTGCTGCTGTTGGT 2949
           |||||| ||||| ||||| || ||| || ||||| ||| |||| |||| ||||
Db   2917 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACACTGCTTCACACTGCTCGGT 2976

Qy   2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTGCTTTA 3009
           |||| ||| ||| ||| ||||| |||| ||||| |||| || ||| |||| ||||| |||
Db   2977 AGTGTGTACAAAAAGATTATTACCCAAACTGGTGTGCTGCGCTGCTAAACCTGCTGCATTG 3036

Qy   3010 GGTGCTGGTAAATCGTGATGATGAAGAAGAACCTGCTGCTGCTGCTCTAAACCTGCTCCAGC 3068
           ||||| ||| ||| ||| |||| |||| ||||| ||||| || |||| ||||
Db   3037 AAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGCTGCTTTTTTTTAAAACCACCTGTTCAACC 3095

RESULT 7
AAT58840
AAT58840_2
Continuation (3 of 6) of AAT58840 from base 200001 (Mycoplasma genitalium genome.)
WP   Sequence split into 6 fragments   LOCUS AAT58840   Accession Aat58840
WP   Fragment Name   Begin   End
WP   AAT58840_0   1   110000
WP   AAT58840_1   100001   210000
WP   AAT58840_2   200001   310000
WP   AAT58840_3   300001   410000
WP   AAT58840_4   400001   510000
WP   AAT58840_5   500001   580073

Query Match   3.9%;   Score 124.6; DB 2;   Length 110000;
Best Local Similarity   63.5%;
Matches 190; Conservative   0; Mismatches 109;   Indels   0; Gaps   0;

Qy   2770 TTTGGGGCCTACCTTCATGGGTAGTGCCTACACGCAATTGGTAGTACATTAGGTATTCTT 2829
           ||||| |||| |||||| |||| | || | |||| |||| || || ||||| |||| |

Db   2765 TTCTTGCAATTATGATCATCTTAGGATTAGCTACGGTATTCCTTTAAGAGCTCAAAGAA 2824
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2885 AATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCTGCTG 2944
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2825 AATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCTGCTG 2884

Qy   2945 TTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTG 3004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2885 TTGGTTCAGTTTACAAGAGATTATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTG 2944

Qy   3005 CTTTAGGTGCTGGTAAATCGGTGATAAGAAGAACCTGCTCGTCGCTGCTAAACCTGCTCTC 3064
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   2945 CTTTAGGTGCTGGTAAATCTGGTGATAAGAAGAAACCTGCTCGTCGCTGCTAAACCTGCTCTC 3004

Qy   3065 CAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGCCTAAATCTG 3124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3005 CAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGCGCCTAAATCTG 3064

Qy   3125 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG 3184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3065 GTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG 3124

Qy   3185 AATAA 3189
           |||||
Db   3125 AATAA 3129

RESULT 6
ACA39446
ID   ACA39446 standard; DNA; 3159 BP.
XX
AC   ACA39446;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #21103.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Mycoplasma genitalium.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,   Zamudio C,   Malone C,   Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,   Trawick JD,   Carr GJ,   Yamamoto R,   Forsyth RA,   Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU35576.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 27316; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid

```
Db  28703 TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG 28762
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2830 GCAATTATGATCATCTTAGGATTAGTTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAAAATTA 2889
                 |||||   ||||||||||||||||||||||||||||||||||||||||||
Db  28763 TTTATCTTGTTAGTCTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 28822
                 ||||||   ||||||||||||||||||||||||||||||||||||||||||
Qy   2890 CAAGACAAAGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
                 |||||   |||||||||||||||||||||||||||||||||||||||||||||
Db  28823 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGTTGATACACTCACCAACTGCTGTCGGT 28882
                 |||||   ||||||||||||||||||||||||||||||||||||||||||
Qy   2950 TCAGTTTACAAGAAGATTATTACCCAAACTGCTACGTTAAGAAAAAACCTGCTGCTTTA 3009
                 |||||   |||||||||||||||||||||||||||||||||||||||||||
Db  28883 AGTGTGTACAAAAAAGATTATTACCCAAACTCGTGTGTGTGAAAAAGCACCTAGTGCATTG 28942
                 ||||   ||||||||||||||||||||||||||||||||||||||||||
Qy   3010 GGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCCTAAACCTGCTGCTCCAGC 3068
                 |||||   |||||||||||||||||||||||||||||||||||
Db  28943 AAAGTCTGCTARATCCTAGTGTGTTAAAAAAACCTGCTGCTTTTTAAAACCACCTGTTCAACC 29001
```

RESULT 8
AAQ81778

ID AAQ81778 standard; DNA; 4182 BP.
XX
AC AAQ81778;
XX
DT 09-AUG-1995 (first entry)
XX
DE DNA encoding cytadhesin protein.
XX
KW Cytadhesin protein; diagnostic probe; unusual codon usage; immunoassay;
KW antigen; ds.
XX
OS Mycoplasma gallisepticum.
XX
FH Key   Location/Qualifiers
FT CDS   713..4081
FT   /*tag= a
FT   /product= "cytadhesin protein"
FT misc_difference 887..889
FT   /*tag= b
FT   /codon= seq:tga, aa:trp
FT misc_difference 998..1000
FT   /*tag= c
FT   /codon= seq:tga, aa:trp
FT misc_difference 1484..1486
FT   /*tag= d
FT   /codon= seq:cta, aa:lys
FT misc_difference 1487..1489
FT   /*tag= e
FT   /codon= seq:tta, aa:lys
FT misc_difference 1559..1561
FT   /*tag= f
FT   /codon= seq:tta, aa:lys
FT misc_difference 1592..1594
FT   /*tag= g
FT   /codon= seq:tta, aa:lys
FT misc_difference 1676..1678
FT   /*tag= h
FT   /codon= seq:tga, aa:trp
FT misc_difference 1715..1717
FT   /*tag= i
FT   /codon= seq:ctt, aa:lys
FT misc_difference 1853..1855
FT   /*tag= j
FT   /codon= seq:tga, aa:trp
FT misc_difference 1890..1892
FT   /*tag= k
FT   /codon= seq:tga, aa:trp
FT misc_difference 1904..1906
FT   /*tag= l
FT   /codon= seq:tta, aa:lys
FT misc_difference 1961..1963
FT   /*tag= m
```

misc_difference 2132..2134
           /*tag= n
           /codon= seq:tga, aa:trp
misc_difference 2237..2239
           /*tag= o
           /codon= seq:ctt, aa:lys
misc_difference 2273..2275
           /*tag= p
           /codon= seq:tga, aa:trp
misc_difference 2417..2419
           /*tag= q
           /codon= seq:tga, aa:trp
misc_difference 2504..2506
           /*tag= r
           /codon= seq:tga, aa:trp
misc_difference 2609..2611
           /*tag= s
           /codon= seq:tta, aa:lys
misc_difference 2639..2641
           /*tag= t
           /codon= seq:tga, aa:trp
misc_difference 2858..2860
           /*tag= u
           /codon= seq:tga, aa:trp
misc_difference 3020..3022
           /*tag= v
           /codon= seq:tga, aa:trp
misc_difference 3083..3085
           /*tag= w
           /codon= seq:tga, aa:trp
misc_difference 3138..3140
           /*tag= x
           /codon= seq:tga, aa:trp
XX
PN  US5378820-A.
XX
PD  03-JAN-1995.
XX
PF  09-NOV-1992;  92US-00973257.
XX
PR  09-NOV-1992;  92US-00973257.
XX
PA  (DOHM/) DOHMS J B.
PA  (KEEL/) KEELER C L.
XX
PI  Keeler CL,  Dohms JE;
XX
DR  WPI; 1995-051314/07.
DR  P-PSDB; AAR64927.
XX
PT  acid encoding cytadhesin protein - used as a probe to diagnose Mycoplasma
PT  gallisepticum infection in poultry.
XX
PS  Claim 1; Col 7-16; 13pp; English.
XX
CC  The sequence encodes a cytadhesin protein from Mycoplasma gallisepticum.
CC  All or part of the sequence may be used as a probe for diagnosis of M.
CC  gallisepticum infection in poultry, e.g. turkey or fowl. The sequence has
CC  several UGA codons, which may be eliminated to avoid problems of
CC  expression in vitro due to irregular codon usage. The sequence may be
CC  cloned in a vector for recombinant cytoadhesin production, for use as an
CC  antigen in immunoassay formats
XX
SQ  Sequence 4182 BP; 1456 A; 763 C; 746 G; 1217 T; 0 U; 0 Other;

Query Match        2.5%;  Score 78.4;  DB 2;  Length 4182;
Best Local Similarity  98.8%;  Pred. No.7.5e-07;
Matches  79;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy     1 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4103 ATGAATATTTCTAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 4162

```
Qy    2907  AACACACATTCAAAAAGTTGATACCTTGACGCTGCTGCTGGTTGGTTGGTTACAAGAAGAT  2966
                  |||||||||||| |||||||     |||||||   |   ||||| |||
Db    3213  ATTGCACACGATAAAGTTGGAACATTGACTTCAGCTGTGGTGGTGTTTTCAAAAAAAT  3272

Qy    2967  TATTACCCAAACTGCTAACGTTAAGAAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGG  3026
                  |||||         ||  ||    |||||       |||  |||||   |||
Db    3273  TATTGACAATACAAATTCTAATAATGTGAAATCTAAACCACAAATGCTTAAGGCTGCAGC  3332

Qy    3027  TGATAAGAAACCTGCTGCTGCTGCTGCTCAAACCTGCTCGTCCAGCTAAACCATCTGCACCAAA  3086
                  |||         |||   ||||    |||      |  ||  ||||| |
Db    3333  TAAAAAACCAAATACAGTTCCACCTGCTAGATCTCAATTAACAAATGATTCTGTTTCTAG  3392

Qy    3087  AGCTAGTCACCAGC  3101
                  ||||| |||||
Db    3393  ACCCACTCCACCATC  3407
```

```
RESULT 10
AAQ75107/c
ID    ABQ75107 standard; cDNA; 4985 BP.
XX
AC    ABQ75107;
XX
DT    01-NOV-2002  (first entry)
XX
DE    Anopheles gambiae gambiae odourant receptor 2 genomic DNA SEQ ID NO:10.
XX
KW    Anopheles gambiae; mosquito; olfactory gene; arrestin 1; pest control;
KW    odourant receptor; olfaction; gene; ds.
XX
OS    Anopheles gambiae.
XX
PN    WO200259274-A2.
XX
PD    01-AUG-2002.
XX
PF    28-JAN-2002; 2002WO-US002549.
XX
PR    26-JAN-2001; 2001US-0264649P.
PR    24-JAN-2002; 2002US-00056405.
XX
PA    (UYVA-) UNIV VANDERBILT.
XX
PI    Zwiebel LJ;
XX
DR    WPI; 2002-627421/67.
DR    P-PSDB; ABP52835.
XX
PT    New mosquito olfaction polypeptides and polynucleotides, useful for
PT    mosquito management, i.e. controlling the pest and disease vectors, or
PT    for identifying pest control agents.
XX
PS    Disclosure; Fig 4a; 96pp; English.
XX
CC    The present invention describes a purified Anopheles gambiae olfaction
CC    polypeptide comprising a 383, 394, 380, 411, 412, 391, 157 or 401 residue
CC    amino acid sequence (see ABP52833 to ABP52840) (S1), a conservatively
CC    modified amino acid sequence of them, or a sequence of (S1) with at least
CC    20 consecutive residues. Also described: (1) an isolated polynucleotide
CC    comprising: (a) a nucleotide sequence encoding the purified Anopheles
CC    gambiae olfaction polypeptide; or (b) a nucleotide sequence that
CC    hybridises under stringent conditions to a hybridisation probe comprising
CC    a 1964,1239, 1142, 1236, 1194, 1176, 474 or 1206 nucleotide sequence
CC    (see ABQ75102 to ABQ75105 and ABQ75110 to ABQ75113) (S2), or its
CC    complement; and (2) a method for identifying an agent that binds to
CC    mosquito olfaction molecule comprising: (a) providing an isolated
CC    mosquito olfaction molecule; (b) contacting a test agent with the
CC    isolated mosquito olfaction molecule; and (c) detecting specific binding
CC    of the test agent to the isolated mosquito olfaction molecule, where the
CC    presence of specific binding identifies the test agent as a mosquito
CC    olfaction-binding compound. The mosquito olfaction molecules are useful
CC    for mosquito management, i.e. controlling this pest and disease vector. A
```

```
Qy      61  GCTCTTGGTTCTCGCAAGCTT  80
                  |||||||||||||||||||
Db    4163  ACTCTTGGTTCTCGCAAGCTT  4182
```

```
RESULT 9
AAQ79746
ID    AAQ79746 standard; DNA; 3435 BP.
XX
AC    AAQ79746;
XX
DT    25-MAR-2003  (revised)
DT    08-SEP-1995  (first entry)
XX
DE    Mycoplasma pirum adhesin gene.
XX
KW    Adhesin; cytoadherence; Mycoplasma; human immunodeficiency virus; HIV;
KW    vaccine; infection; ss.
XX
OS    Mycoplasma pirum.
XX
FH    Key          Location/Qualifiers
FT    CDS          1..3435
FT                 /tag= a
FT                 /codon= TGA; aa:Trp
FT                 /note= "typical of Mycoplasma"
XX
PN    FR2705970-A1.
XX
PD    09-DEC-1994.
XX
PF    04-JUN-1993; 93FR-00006745.
XX
PR    04-JUN-1993; 93FR-00006745.
XX
PA    (INSP ) INST PASTEUR.
XX
PI    Tham TN, Ferris S, Blanchard A, Montagnier L, Bahraoui E;
XX
DR    WPI; 1995-024735/04.
DR    P-PSDB; AAR76059.
XX
PT    Nucleic acid encoding Mycoplasma pirum adhesin - and related proteins and
PT    antibodies, useful in diagnosis, treatment and prevention of M.pirum and
PT    HIV infection.
XX
PS    Claim 2; Fig 1; 55pp; French.
XX
CC    The M.pirum adhesin gene was isolated using a probe based on the 3'- end
CC    of the known M.genitalium adhesin gene. M.pirum is found in AIDS patients
CC    and the adhesin may have a role in infection of cells by HIV. If so, the
CC    M.pirum adhesin protein will be useful in vaccines to protect against the
CC    cytopathic effect of the Mycoplasma and against that of HIV. (Updated on
CC    25-MAR-2003 to correct PN field.)
XX
SQ    Sequence 3435 BP; 1334 A; 466 C; 494 G; 1141 T; 0 U; 0 Other;
```

```
Query Match      2.0%;  Score 64.6;  DB 2;  Length 3435;
Best Local Similarity  48.3%;  Pred. No. 0.00068;
Matches  181;  Conservative  0;  Mismatches 194;  Indels  0;  Gaps  0;

Qy    2727  TRAATTAAGAACAATAATTAGTAACATTCCTTATAACAACTTTCGCGCCTTACCTTC  2786
                  ||| |||      ||   |||||    |   ||  ||||||   |||
Db    3033  TAGTTCTAAAATAAATGTTATAAAATAATTCTATTTTTGCTGGATTTAGTGCAATGGCTGA  3092

Qy    2787  ATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTTGCAATTATGATCATCTT  2846
                  ||| ||||||     | ||  ||   || |||     |||     ||||||
Db    3093  TTGAATTCTTCCTGTAGTTATTGCTATTCCCATTGTATTAGTTGCATTAAATAATTGGTTT  3152

Qy    2847  AGGATTAGCTATCGTATTCCTTAAGAGCTCAAAGAAAATTACAAGACAAAGGGTTCAA  2906
                  |||||||||||||||   |||       || ||||  || |||
Db    3153  AGGTTGCAGTATTGGGATTCCAATGCCTAAACATAAAAACATAAAAAAGCTATTAAAGTTGGTTTTGA  3212
```

CC   method from the present invention of screening for substances that
CC   modulate arrestin-odourant receptor interaction is useful for identifying
CC   pest control agents. The present sequence represents Anopheles gambiae
CC   odourant receptor 2 genomic DNA from the present invention. N.B. The
CC   features given in figure 4a are tentative and do not directly encode SEQ
CC   ID NO:6 (ABP52835)
XX
SQ   Sequence 4985 BP; 1319 A; 1058 C; 990 G; 1618 T; 0 U; 0 Other;

Query Match   2.0%; Score 62.6; DB 6; Length 4985;
Best Local Similarity 45.5%; Pred. No. 0.0021;
Matches 333; Conservative   0; Mismatches 396; Indels   3; Gaps   3;

Qy   253 ATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGTCAAGTGATGATTTGGT 312
           ||| | || |||| ||| || ||| | | |||| | ||| || || || |||
Db   3070 ATAAAAAAACAATAATAAACATAATAATAAGAATAACAACAACAATAATATAAGAA 3011

Qy   313 GGATACAAAGTAAAACAAAATAGTTTCA-GATTACCAACTAGCAGAAATAGATTTGATCA 371
           |||| || | || || || ||| | | | | || ||
Db   3010 CAACAACAACAATAATAAGAATATATAATAATAACAATAATAAGAATAATAA 2951

Qy   372 AAGACAAACAAGAGCATATTATGCTCGTTGGTAATGATGAAGCTAACGTTCAT-TTAA 430
           || | | | |||||| || || | | || || || ||| || || ||
Db   2950 TAACAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAA 2891

Qy   431 AAAGAATTAATACTAACTCAAATAGAATTGGTAAATAGAAACAACAATTCTAAGTTGTAA 490
           ||| ||| || | ||| |||| ||| || || | | |||| |||| ||||
Db   2890 AAAAAATAATTATATATAAGAATAGTAATAATAATAATAATAATTAAAAAATAATATTA 2831

Qy   491 TTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAATTTA 550
           | | | | | ||||| | || |||| || || | |||| | || |||||
Db   2830 ATAATAATAATAATAATAATAATAAGAATAATAATAATAATAATAATAATA 2771

Qy   551 ATTTTACAAACCAAACTCAAGGTGAAATTGTAAGTGACTTCATTTTAGATGCGCCAATCT 610
           || || | |||| ||| | |||| | | | || || || | | || ||||
Db   2770 ATAATAATAATAATCATATAAAAATTATTATTGATAATAATGTATAATAATGAAT 2711

Qy   611 TAC-CTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAGAGATCTTACCA 669
           ||| |||| |||| ||| | | ||| || | | | ||| || | |||| | ||
Db   2710 AACAATAATAATAATAATTAATAAATAATATAATTAATTAATAATATAATATAATAAT 2651

Qy   670 AATGACGTCAACACTGCAGTTGTTCCTTGCCAGTAGGTAGTTAGTGGAACAAATGCT 729
           ||||| |||| | || ||| |||| || || || || || | |||||| ||
Db   2650 TATTATTGTTTATAATAATAATAATAATAATAATCATAATCATAATTATAT 2591

Qy   730 GATGATGGGGATGTTTGATTGTGGGAAGTGTCAAATAACTAATACAGATCCTATTGCTCAA 789
           || ||| ||| | | || ||| ||| || | | ||| ||| | | | || ||| ||
Db   2590 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAA 2531

Qy   790 ACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGCCTGGT 849
           ||| ||| | || || ||||| |||| | || |||||| ||||| || | ||
Db   2530 AGTAATAATTATATATAGTAATAATAATAATAATAATATTAATAATAAT 2471

Qy   850 GCAAACAAATAGATACGATTCTCAATTGAAATGTCAAGCATAGAATTAAAACATCTTCCAA 909
           || ||| ||| | ||| || ||| || ||| |||| || ||| | |||| ||| ||||
Db   2470 AATAATAATAATAATAATAATTAATAATAATTATAATAATAATAATAAT 2411

Qy   910 TTAGATGAAAAATTTGTTTATCCGAAATGGACTGGTTCTGAAGAGAATAAAAAATATTACA 969
           || ||| ||||| | | ||| | ||| | | ||| || || |||||| ||| ||||
Db   2410 AATAATAATAATAATAATAATAATAATAATAATAATATAATAATAATAATAAT 2351

Qy   970 AGATTAGCTACT 981
           | |||| | |
Db   2350 AAATTCTGTATT 2339

RESULT 11
ACF79720/c
ID   ACF79720 standard; DNA; 4985 BP.
XX
AC   ACF79720;
XX
DT   15-JAN-2004 (first entry)

XX
DE   Mosquito odorant receptor 2 genomic DNA.
XX
KW   Mosquito; odorant receptor 2; olfaction; insecticide; antimalarial; gene;
KW   ss.
OS   Anopheles gambiae.
XX
FH   Key   Location/Qualifiers
FT   exon   1..702
FT   /*tag= d
FT   /number=1
FT   5'UTR   1..485
FT   /*tag= a
FT   TATA_signal   295..300
FT   /*tag= b
FT   CDS   486..3561
FT   /*tag= c
FT   /product= "Odorant receptor 2"
FT   /note="contains introns"
FT   intron   703..798
FT   /*tag= e
FT   /number=1
FT   exon   799..882
FT   /*tag= f
FT   /number=2
FT   intron   883..978
FT   /*tag= g
FT   /number=2
FT   exon   979..1404
FT   /*tag= h
FT   /number=3
FT   intron   1405..1601
FT   /*tag= i
FT   /number=3
FT   exon   1602..1702
FT   /*tag= j
FT   /number=4
FT   intron   1703..1783
FT   /*tag= k
FT   /number=4
FT   exon   1784..1886
FT   /*tag= l
FT   /number=5
FT   intron   1887..1993
FT   /*tag= m
FT   /number=5
FT   exon   1994..2106
FT   /*tag= n
FT   /number=6
FT   intron   2107..3459
FT   /*tag= o
FT   /number=6
FT   exon   3460..4985
FT   /*tag= p
FT   3'UTR   3562..4985
FT   /*tag= q
XX
PN   WO2003076590-A2.
XX
PD   18-SEP-2003.
XX
PF   10-MAR-2003; 2003WO-US007174.
XX
PR   08-MAR-2002; 2002US-00094240.
XX
PA   (UTVA-) UNIV VANDERBILT.
XX
PI   Zwiebel LJ;
XX
DR   WPI; 2003-722331/68.
XX
PT   New mosquito arrestin 1 and 2 genes and polypeptides, useful for

PT   identifying mosquito olfaction molecule binding compounds which reduce
PT   the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
XX
PS   Disclosure; Fig 4a; 101pp; English.
XX
CC   The present sequence is that of genomic DNA encoding a novel mosquito
CC   olfaction molecule, odorant receptor 2. A cDNA sequence for odorant
CC   receptor 2 is given in ACF79719. The invention provides 9 novel mosquito
CC   polypeptides and the nucleic acids encoding them. These are odorant
CC   receptor molecules 1-7 and arrestins 1 and 2. The odorant receptors
CC   function in a ligand-induced signal transduction pathway for the
CC   activation of mosquito olfaction. Arrestin functions to inhibit the
CC   activated signal transduction cascade. Thus, the odorant receptors act as
CC   an 'on' switch, and arrestin as an 'off' switch for the odorant detection
CC   system of the mosquito. Methods are provided for identifying compounds
CC   that interfere with the operation of the mosquito olfactory system,
CC   particularly compounds that modulate arrestin 2 activity. These are
CC   useful for the control of mosquitoes, particularly by reducing their
CC   ability to locate sources of bloodmeal
XX
SQ   Sequence 4985 BP; 1319 A; 1058 C; 990 G; 1618 T; 0 U; 0 Other;

Query Match   2.0%; Score 62.6; DB 9; Length 4985;
Best Local Similarity 45.5%;   Pred. No. 0.0021;
Matches 333; Conservative   0; Mismatches 396; Indels   3; Gaps   3;

Qy   253 ATCAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGTCATGATGATTGTCGT 312
            ||||| || || ||| ||| ||| | | | || || ||| |||| | || ||| |
Db   3070 ATAAAAAAAAAACATTATAATAAACAATATAACAACAACAACAATAATAAGAA 3011

Qy   313 GGATACAAAGTAAAACAAATAGTTTCA-GATTACACAACTAGCACAGAAATAGATTTGATCA 371
            |||| | | | || || || ||| ||| |||||| | ||| || |||| | |||
Db   3010 CAACACACAACAACAATAATAAGAATAATAATAACAATAGCAATAATAAGAATAATAA 2951

Qy   372 AAGACAAAACAAGAGCGATATTATGTCTCTGTTGGTTAATGAGAAGCTAACGTTCAT-TTAA 430
            || | | ||| | || ||| ||| | | ||| || || | ||| ||||
Db   2950 TAACAATAGCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 2891

Qy   431 AAAGAATTAATACTAACTCAAATAGATGGTATAGAAACAACAACAATTCTAAGTTTGTAA 490
            || | || ||| || ||| | ||| | | | ||| | ||||| ||| | | |
Db   2890 AAAAAAATAATTAATAATAAGAATAGTAATAATAATAATAATAATAAAAATATAATTA 2831

Qy   491 TTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAAATTA 550
            || | | | ||| || || | | | | || || ||| | || | || | | |||| ||
Db   2830 ATAATAATAATAATAATAATAATAAGAATAATAATAATAATAATAATAATAATAATA 2771

Qy   551 ATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCT 610
            || ||||| || | ||| | || ||| || | | || ||| ||| | | | || |
Db   2770 ATAATAATAATAATCATAATAAATTATTATTATGATAATAATGATATAATAATGAAT 2711

Qy   611 TAC-CTAAAGATTTACACCCAGTTGGTATACCTTATAACTTATAACATTCAAAGAAGATCTTACCA 669
            ||| |||| || || | | ||| || ||| | | || |||| ||||| |||||| ||||||
Db   2710 AACAATAATAATAATTAATAAATTAATAAATAATAATAATAATAATAATAATAATAAT 2651

Qy   670 AATGACGTCAACACTGCAGTTGTTCCTTGGCCCAGTAGGTTAGTGGAACAAAATGCT 729
            ||| || || || | | | || || | | |||| | || | | ||| |||
Db   2650 TATTATTGTTATATAATAATAATAATAATAATAATCATAATTCATAATTATAAT 2591

Qy   730 GATGATGGGGATGGTTGATTGTTGGGAATGGTCAAATAACTAATACAGATCCTAATGCTCAA 789
            |||||||| |||| ||||| | |||||||| ||||| | || ||| | || ||| || || |
Db   2590 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAAATAATAAA 2531

Qy   790 ACTAAAACCACTACTGATAATCAAAATCCTCAACTTTTAATTCAGGAGCAATGCCTCGT 849
            ||| | | || || ||| ||| | || | | ||| |||||| |||| ||| | | |
Db   2530 AGTAATAATAATAATAGTAATAATAATAATAATAATAATAATAATAATAATAATAAT 2471

Qy   850 GCAAACAATAGATACGATTCTCATTGAATGTCAAGCATAGAATTAAAACATCTTTCCAA 909
            |||||||| | || | ||| |||| || | || || | || ||||| || || |||
Db   2470 AATAATAATAATAATAATTATTATTAATTATAATAATAATAATAATAATAATAATAAT 2411

Qy   910 TTAGATGAAAAATTTGTTTATCCAGAATGGACTCGTTCTGAAGAGAATTAAAAATATTACA 969
            ||||| | ||| | ||| ||||| || || | ||| | ||||| |||||||| ||
Db   2410 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 2351

Qy   970 AGATTAGCTACT 981
            |||||||| |||
Db   2350 AAATTCTGTATT 2339

RESULT 12
ABL33825/c
ID   ABL33825 standard; DNA; 8771 BP.
XX
AC   ABL33825;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1798.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A, Piepenbrock C, Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 1798; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 8771 BP; 2734 A; 59 C; 1660 G; 4318 T; 0 U; 0 Other;

Query Match   1.9%; Score 60.8; DB 6; Length 8771;
Best Local Similarity 46.3%;   Pred. No. 0.006;
Matches 200; Conservative   0; Mismatches 232; Indels   0; Gaps   0;

Qy   256 AAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGTCATGATGATTGTGGTGGA 315
            |||||||||| |||| ||||| |||||| ||||||| |||||| |||| || || |||
Db   5053 ATAATCTATACATACAATAAATTATTCAACTTTAAAAAAAAAAAAACTTAATCTA 4994

Qy   316 TACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCACAACTAGCAGAAATAGATTTGATCAAGA 375
            ||| | ||| |||||| || | ||| | | | || || | | | ||| |||||
Db   4993 TACTACACAAAAAAAAAAACCTTAAAAACATTAAAAATAAAATAAAACCAATAACAAAAA 4934

Qy   376 CAACAAGAGCACATATATGCTCGTGTTGGTTAATGATGATGATGAAGCTAACGTTCATTTAAAAAGA 435
            |||

Db   4933 AATACATACTATATAAATTCCACTTCTAAAAACTACCTAAAAATCAAAATTATAAAAAC 4874

Qy   436 ATTAATAACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGT 495
            | ||||| ||| || | |||| | | || ||| || || |||||| ||| |||
Db   4873 AAAAATAATAACATAAATTCAAAAACCTAAAAAAAAAAAATAAAAAAT'TATTTATTTAAT 4814

Qy   496 GGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGCATGATGGGACTAAATTAATTTT 555
            ||| | || ||| | | || | | ||| | |||||||
Db   4813 AAATAATAATAATATACAAAAATATAAACATAATTAATACCCACTAAATTATACACA 4754

Qy   556 ACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGATGCGCCAATCTTACCT 615
            || || | || | | | || ||| || | || || | | | ||||||
Db   4753 TACAAATAATCAAAATAATAAATTTTATATTATATCATTCAAAGATCTTACATATAATTTTACCA 4694

Qy   616 AAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCTTACCAAATGAC 675
            || | || | | | | | || || | | ||| | || | |
Db   4693 TAATACAAAAATAAAAAAAAAAAAAAAATAAAAAACTTAACGATTTACATAACTTA 4634

Qy   676 GTCAACACTGCA 687
            || || ||||
Db   4633 ACCAATACCTCA 4622

RESULT 13
AAA30290/c
ID   AAA30290 standard; DNA; 3489 BP.
XX
AC   AAA30290;
XX
DT   15-SEP-2003 (revised)
DT   11-SEP-2000 (first entry)
XX
DE   Kaposi's sarcoma-associated herpesvirus LANA gene.
XX
KW   Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
KW   latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
KW   Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
KW   Kaposi's sarcoma; primary effusion lymphoma; PEL;
KW   human immunodeficiency virus; HIV; multicentric Castleman's disease; ds.
XX
OS   Human herpesvirus 8.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..3489
FT   /*tag= a
FT   /product= "LANA"
FT   misc_signal   40..50
FT   /*tag= b
FT   /note= "nuclear localisation signal, NLS"
FT   misc_signal   190..210
FT   /*tag= c
FT   /note= "nuclear localisation signal, NLS"
XX
PN   WO200029625-A1.
XX
PD   25-MAY-2000.
XX
PF   19-NOV-1999; 99WO-US027508.
XX
PR   19-NOV-1998; 98US-00109422.
PR   21-APR-1999; 99US-0298568.
XX
PA   (KIEF/) KIEFF E D.
PA   (BALL/) BALLESTAS M E.
PA   (KAYE/) KAYE K M.
XX
PI   Kieff ED, Ballestas ME, Kaye KM;
XX
DR   WPI; 2000-387829/33.
DR   P-PSDB; AAV96255.
XX
PT   Treating or preventing a disease associated with rhodino virus infection
PT   in a mammal which includes Kaposi's Sarcoma and Primary Effusion

PT   Lymphoma.
XX
PS   Disclosure; Fig 6; 70pp; English.
XX
CC   The present sequence is the Kaposi's sarcoma-associated herpesvirus
CC   (KSHV) latency-associated nuclear antigen (LANA) gene. KSHV is also known
CC   as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma
CC   -2 herpes virus class. The LANA protein is necessary for the efficient
CC   persistence of rhadino virus DNA in mammalian cells. Persistent rhadino
CC   virus infection is implicated in a variety of diseases e.g. Kaposi's
CC   Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric
CC   Castleman's disease. In addition, KS is a common malignancy in HIV
CC   patients. KSHV persists in host cells in a latent form. One of the few
CC   genes expressed from the latent viral DNA is LANA. LANA associates with
CC   both human chromosomes and with the rhadino virus cis-acting element
CC   (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
CC   "tied" to the host chromosomes. This allows the viral DNA to persist in
CC   the host cell. The present sequence may be used to screen and identify
CC   molecules that inhibit LANA interaction with RVCAE, thereby interfering
CC   with the latency cycle of this virus. Potential antiviral treatments for
CC   the above mentioned diseases may therefore be based on LANA deregulation.
CC   (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ   Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;
XX

Query Match   1.9%;   Score 60.2; DB 3;   Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0062;
Matches 114; Conservative 0; Mismatches 68; Indels 3; Gaps   1;

Qy   2998 CCTGCTGCTTAGGTGCTGCTGGTGAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT 3057
            ||||||||||||||| ||| ||| || ||| |||| | ||||||||||||||||||
Db   2236 CATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCT 2177

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
            |||||| ||| || | ||||||||| | |||| ||||| || ||| ||||
Db   2176 GCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCCT 2117

Qy   3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTGCTCCTAGCCAGCTCAGCTCCAAAACCAACC 3174
            |||||| | | || || || | |||| ||||| |||| || | |||| | | ||
Db   2116 GCTCATCCTGCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCT 2057

Qy   3175 GCTCC 3179
            |||||
Db   2056 GCTGC 2052

RESULT 14
AAF82901/c
ID   AAF82901 standard; DNA; 3489 BP.
XX
AC   AAF82901;
XX
DT   11-SEP-2003 (revised)
DT   06-AUG-2003 (revised)
DT   29-JUN-2001 (first entry)
XX
DE   Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW   Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW   Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW   KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS   Human herpesvirus 8.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..3489
FT   /*tag= a
XX
PN   WO200125484-A2.
XX
PD   12-APR-2001.
XX
PF   29-SEP-2000; 2000WO-US026908.

XX
PR 01-OCT-1999; 99US-00410399.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Robertson ES, Cotter MA;
XX
DR WPI; 2001-281736/29.
DR P-PSDB; AAB62331.
XX
PT A composition for use in gene therapy comprises an expression vector that includes a nucleic acid sequence encoding a nucleic acid binding protein.
XX
PS Disclosure; Fig 9A; 60pp; English.
XX
CC The invention provides a composition comprising nucleic acid, histone H1 protein and expression vector operationally encoding a protein suitable for tethering the nucleic acid to the histone H1 protein, where the tethering protein is LANA. The composition is useful in aiding the retention of the viral DNA in the host cell. The viral vector encodes a protein suitable for tethering DNA to Histone H1. Methods for screening for compounds which are agonistic or antagonistic for the tethering of viral proteins to histone H1 and DNA binding sites are useful for developing the method of viral transfer. The composition has applications to gene therapy, including the treatment of multiple sclerosis, Parkinson's disease, Huntington disease and diabetes. The present sequence represents the nucleotide sequence of the Kaposi's sarcoma associated herpesvirus (human herpesvirus 8) latency-associated nuclear antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-2003 to correct OS field.) (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

```
Query Match             1.9%;  Score 60.2;  DB 4;  Length 3489;
Best Local Similarity  61.6%;  Pred. No. 0.0062;
Matches  114;  Conservative   0;  Mismatches  68;  Indels   3;  Gaps    1;

Qy   2998 CCTGCTGCTTAGGTGCTGTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT 3057
               |||||||||||||||||||||||||||
Db   2236 CATCCTGCTGCTGCTGTCCATCCTGCTGCTGCTGCTGCTCATCCT 2177

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGC-- 3115
Db   2176 GCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCCT 2117

Qy   3116 -CTAAATCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
Db   2116 GCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCTGCT 2057

Qy   3175 GCTCC 3179
Db   2056 GCTGC 2052
```

RESULT 15
ABA93487/c
ID ABA93487 standard; DNA; 3489 BP.
XX
AC ABA93487;
XX
DT 25-APR-2002 (first entry)
XX
DE Kaposi's sarcoma-associated herpesvirus LANA protein encoding DNA.
XX
KW Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;
KW KSHV terminal repeat; rhadino virus cis acting element; episome;
KW primary effusion lymphoma; latency-associated nuclear antigen;
KW gene therapy; gene transfer; gene; ds.
XX
OS Human herpesvirus 8.
XX
FH Key   Location/Qualifiers

FT CDS   1..3489
FT   /*tag= a
FT   /product= "LANA protein"
FT   /note= "latency-associated nuclear antigen"
XX
PN US6322792-B1.
XX
PD 27-NOV-2001.
XX
PF 21-APR-1999; 99US-0029855B.
XX
PR 19-NOV-1998; 98US-0109422P.
XX
PA (KIBF/) KIEFF E D.
XX
PI Kieff ED, Ballestas ME, Kaye KM;
XX
DR WPI; 2002-153769/20.
DR P-PSDB; ABB05621.
XX
PT System for episomal retention of plasmids in mammalian cells, useful in gene therapy, comprises rhadinoviral LANA and RVCAE sequences.
XX
PS Claim 1; Fig 6; 27pp; English.
XX
CC The present invention describes a system (A) for maintaining a plasmid as an episome in mammalian cells, comprising the rhadinoviral sequence LANA (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487, S1) expressed in the cell, and the rhadinoviral sequence RVCAE (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2) present in the plasmid. Also describes a method for maintaining a closed circular DNA in a cell by expressing (S1) in the cells and having (S2) as a cis-acting and maintenance sequence in the DNA. (A) is particularly used in gene therapy (or other gene transfer applications) that uses mammalian cells in which LANA is expressed. (A) improves persistence of gene therapy vectors in cells. The present sequence encodes Kaposi's sarcoma-associated herpesvirus (KSHV, also called human herpesvirus 8) LANA protein, which is used in the exemplification of the present invention
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;

```
Query Match             1.9%;  Score 60.2;  DB 6;  Length 3489;
Best Local Similarity  61.6%;  Pred. No. 0.0062;
Matches  114;  Conservative   0;  Mismatches  68;  Indels   3;  Gaps    1;

Qy   2998 CCTGCTGCTTAGGTGCTGTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT 3057
               |||||||||||||||||||||||||||
Db   2236 CATCCTGCTGCTGCTGTCCATCCTGCTGCTGCTGCTGCTCATCCT 2177

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCCACCAGCTAAACCAACTGGGC-- 3115
Db   2176 GCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCCT 2117

Qy   3116 -CTAAATCTGGTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
Db   2116 GCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCTGCTGCTCTGCT 2057

Qy   3175 GCTCC 3179
Db   2056 GCTGC 2052
```

Search completed: May 5, 2004, 12:25:53
Job time : 1125.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2004, 10:52:31 ; Search time 240.505 Seconds
                                       (without alignments)
                                       7358.420 Million cell updates/sec

Title: US-09-901-572A-2
Perfect score: 3189
Sequence: 1 atgaatattctaaaaact.........caaccgctcccaagaataa 3189

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
              1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
              6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

              Pred. No. is the number of results predicted by chance to have a
              score greater than or equal to the score of the result being printed,
              and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	124.6	3.9	580073	4	US-08-545-528D-1	Sequence 1, Appli
2	78.4	2.5	4182	4	US-07-973-257-1	Sequence 1, Appli
3	60.2	1.9	3489	2	US-08-728-323A-1	Sequence 1, Appli
c 4	60.2	1.9	3489	4	US-09-298-568-1	Sequence 1, Appli
c 5	60.2	1.9	3489	4	US-09-410-399-1	Sequence 1, Appli
6	60.2	1.9	32207	2	US-08-770-379-20	Sequence 20, Appl
7	60.2	1.9	32207	4	US-08-757-669A-20	Sequence 20, Appl
8	60.2	1.9	32207	4	US-09-230-371A-20	Sequence 20, Appl
9	58.4	1.8	640681	4	US-09-790-988-1	Sequence 1, Appli
c 10	54.2	1.7	832	4	US-09-621-976-2813	Sequence 2813, Ap
11	53.2	1.7	987	4	US-09-142-584-5	Sequence 5, Appli
12	53.2	1.7	1037	4	US-09-181-585-3	Sequence 3, Appli
13	53.2	1.7	1159	4	US-09-181-585-1	Sequence 1, Appli
14	53.2	1.7	1471	4	US-09-181-585-2	Sequence 2, Appli
15	53	1.7	987	4	US-09-142-584-3	Sequence 3, Appli
16	52.4	1.6	3057	4	US-09-601-198-55	Sequence 55, Appli
17	51.4	1.6	987	4	US-09-142-584-1	Sequence 1, Appli
c 18	51	1.6	636	3	US-08-998-416-1137	Sequence 1137, Ap
19	51	1.6	1575	4	US-09-543-681A-615	Sequence 615, App
20	49.8	1.6	300	4	US-09-328-352-4065	Sequence 4065, Ap
c 21	49.4	1.5	1887	4	US-09-601-198-39	Sequence 39, Appl
22	49	1.5	5163	3	US-08-700-651-1	Sequence 1, Appli
23	49	1.5	5163	3	US-08-928-361B-4	Sequence 4, Appli
24	49	1.5	5163	3	US-09-588-995A-4	Sequence 4, Appli
25	49	1.5	5318	3	US-08-700-651-2	Sequence 2, Appli
26	49	1.5	5318	3	US-08-928-361B-3	Sequence 3, Appli
27	49	1.5	5318	4	US-09-588-995A-3	Sequence 3, Appli

c 28	48.4	1.5	397	3	US-09-253-691-3	Sequence 3, Appli
29	48.4	1.5	4884	1	US-07-665-792B-10	Sequence 10, Appli
c 30	48.2	1.5	234	1	US-08-469-802B-3	Sequence 3, Appli
c 31	48.2	1.5	234	2	US-08-267-803B-3	Sequence 3, Appli
c 32	47.6	1.5	195	1	US-08-469-802B-2	Sequence 2, Appli
c 33	47.6	1.5	195	2	US-08-267-803B-2	Sequence 2, Appli
c 34	47.4	1.5	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c 35	47.2	1.5	165	3	US-09-043-303-17	Sequence 17, Appl
c 36	47.2	1.5	203	3	US-09-043-303-7	Sequence 7, Appl
c 37	47.2	1.5	429	3	US-09-543-681A-707	Sequence 707, App
38	46.2	1.4	5511	3	US-08-928-361B-2	Sequence 2, Appli
39	46.2	1.4	5511	4	US-09-588-995A-2	Sequence 2, Appli
40	46.2	1.4	7334	3	US-08-928-361B-1	Sequence 1, Appli
41	46.2	1.4	7334	4	US-09-588-995A-1	Sequence 1, Appli
42	46	1.4	1588	4	US-09-490-291-7	Sequence 7, Appli
c 43	46	1.4	5181	3	US-08-257-073-10	Sequence 10, Appli
44	46	1.4	16442	3	US-08-781-891-208	Sequence 208, App
45	46	1.4	16442	4	US-09-618-166-208	Sequence 208, App

ALIGNMENTS

RESULT 1
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
;   APPLICANT: Fraser et al.
;   TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmen
;   Patent No. 6537773   TITLE OF INVENTION: Thereof, and Uses Thereof
;   FILE REFERENCE: PB193P1
;   CURRENT APPLICATION NUMBER: US/08/545,528D
;   CURRENT FILING DATE: 1995-10-19
;   PRIOR APPLICATION NUMBER: US 08/488,018
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: US 08/473,545
;   PRIOR FILING DATE: 1995-06-07
;   NUMBER OF SEQ ID NOS: 1
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 580073
;   TYPE: DNA
;   ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

        Query Match   3.9%; Score 124.6; DB 4; Length 580073;
        Best Local Similarity 63.5%; Pred. No. 4e-19;
        Matches 190; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy   2770 TTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTATTCTT 2829
              |||||
Db   228703 TTTGCAGCACTCCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG 228762

Qy   2830 GCAATTATGATCATCTTAGGATTAGCTAGCTGTCATCGGTATCTTTAAGAGCTCAAAGAAATTA 2889
              |||||
Db   228763 TTTATCTTGTTAGTCTTTAGGACTTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC 228822

Qy   2890 CAAGACAAAGGGTTCAAAACAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTGTTGGT 2949
              |||||
Db   228823 CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACATCCTTGCTGTCGTGTCGGT 228882

Qy   2950 TCAGTTTACAAGAGAGATTATTATACCCAAACTGCTAAACTGTTAAGAAAAAAACCTGCTGCTTTA 3009
              |||||
Db   228883 AGTGTGTACAAAAGATTATTATACCCAAACTGGTGTGTGGTGAAAAAAGCACCTAGTGCATTG 228942

Qy   3010 GGTGCTGGTAAATCTGGTGATAAGAAACCTGCTCTGCTGTGCTAAACCTGTCGTCCAGC 3068
              |||||
Db   228943 AAAGCTGCTAATCCTAGTGTTAAAAAACCTGCTGCTTTTTAAAACCACCTGTTCAACC 229001

RESULT 2

US-07-973-257-1
; Sequence 1, Application US/07973257
; Patent No. 5378820
; GENERAL INFORMATION:
; APPLICANT: Keeler, Jr. Calvin L.
; APPLICANT: Dohms, John B.
; TITLE OF INVENTION: Gene Encoding Cytadhesin
; TITLE OF INVENTION: Protein of Mycoplasma Gallisepticum and Its Use
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Connolly and Hutz
;   STREET: 1220 Market Street
;   CITY: Wilmington
;   STATE: Delaware
;   COUNTRY: U.S.A.
;   ZIP: 19899
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM/PC or Compatible
;   OPERATING SYSTEM: MS-DOS
;   SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/973,257
;   FILING DATE: 19921109
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA: No. 5378820e
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4182 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: double stranded
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
;   HYPOTHETICAL: No
;   ANTI-SENSE: No
;   ORIGINAL SOURCE:
;   ORGANISM: Mycoplasma Gallisepticum
;   STRAIN: S6
US-07-973-257-1

Query Match   2.5%; Score 78.4; DB 1; Length 4182;
Best Local Similarity 98.8%; Pred. No. 5e-09;
Matches   79; Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4103 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGA 4162

Qy   61 GCTCTTGGTTCTGCAAGCTT 80
              ||||||||||||||||||||
Db   4163 ACTCTTGGTTCTGCAAGCTT 4182

RESULT 3
US-08-728-323A-1/c
; Sequence 1, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Cooper & Dunham LLP
;   STREET: 1185 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.

;   ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/728,323A
;   FILING DATE:
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: White, John P.
;   REGISTRATION NUMBER: 28,678
;   REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 212-278-0400
;   TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3489 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..3489
US-08-728-323A-1

Query Match   1.9%; Score 60.2; DB 2; Length 3489;
Best Local Similarity 61.6%; Pred. No. 0.0001;
Matches   114; Conservative   0; Mismatches 68; Indels   3; Gaps   1;

Qy   2998 CCTGCTGCTTAGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
              ||||||| |||||||| |||||||||||||||||| |||||||||||||||||||
Db   2236 CATCCTGCTGCTGCTGCTCATCCTGCTGCTGCTCATCCTGCTGCTGCTGCTCATCCT 2177

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2176 GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTGCTCCTGCTCCT 2117

Qy   3116 -CTAAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCCAGCGCTCCAAAACCAACC 3174
               |||||||||||||||||||||||||||||||||||||||||
Db   2116 GCTCATCCTGCTGCTGCTCATCCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy   3175 GCTCC 3179
              ||||
Db   2056 GCTGC 2052

RESULT 4
US-09-298-568-1/c
; Sequence 1, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; CURRENT FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; EARLIER FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

```
Query Match        1.9%;   Score 60.2;  DB 4;  Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0001;
Matches 114;  Conservative    0;  Mismatches  68;  Indels   3;  Gaps    1;

Qy   2998  CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
                 |||||  ||||||||||||||||||||||||||||||||||||||||| |||||
Db   2236  CATCCTGCTGCTGCTGCTCATCCTGCTCATCCTGCTCGTGCTGCTGCTCATCCT 2177

Qy   3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCGCTAAACCAACTGGGC-- 3115
                 ||||||| |||||||  |||||||||||||||||||||||||||||||||||||
Db   2176  GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTCATCCTGCTGCTGCTGCTCCT 2117

Qy   3116  -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
                 |||||||| |||  |||||||||||||| |||||||||||||||| |||||
Db   2116  GCTCATCCTGCTCCTGCTCCTCGTGCTCCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy   3175  GCTCCC 3179
                 |||||
Db   2056  GCTGC 2052

RESULT 5
US-09-410-399-1/c
; Sequence 1, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
;  APPLICANT:  Robertson, Erle S.
;  TITLE OF INVENTION:  Methods to Inhibit or Enhance the Binding of Viral DNA
;  TITLE OF INVENTION:   to Genomic Host DNA
;  FILE REFERENCE:  UM-03778
;  CURRENT APPLICATION NUMBER:  US/09/410,399
;  CURRENT FILING DATE:  1999-10-01
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 1
;  LENGTH:  3489
;  TYPE:  DNA
;  ORGANISM:  Kaposi's sarcoma-associated herpesvirus
US-09-410-399-1

Query Match        1.9%;   Score 60.2;  DB 4;  Length 3489;
Best Local Similarity   61.6%;   Pred. No. 0.0001;
Matches 114;  Conservative    0;  Mismatches  68;  Indels   3;  Gaps    1;

Qy   2998  CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
                 |||||  ||||
Db   2236  CATCCTGCTGCTGCTGCTCATCCTGCTCATCCTGCTCGTGCTGCTGCTCATCCT 2177

Qy   3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCGCTAAACCAACTGGGC-- 3115
                 ||||||| |||||||  |||||||
Db   2176  GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTCATCCTGCTGCTGCTGCTCCT 2117

Qy   3116  -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
                 |||||||| |||  ||
Db   2116  GCTCATCCTGCTCCTGCTCCTCGTGCTCCTGCTCATCCTGCTGCTCATCCTGCT 2057

Qy   3175  GCTCCC 3179
                 |||||
Db   2056  GCTGC 2052

RESULT 6
US-08-770-379-20
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
;  APPLICANT:  Chang, Yuan
;  APPLICANT:  Bohenzky, Roy A.
;  APPLICANT:  Russo, James J.
;  APPLICANT:  Edelman, Isidore S.
;  APPLICANT:  Moore, Patrick S.
;  TITLE OF INVENTION:  POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
```

```
TITLE OF INVENTION:  HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
NUMBER OF SEQUENCES:  20
CORRESPONDENCE ADDRESS:
  ADDRESSEE:  Cooper & Dunham LLP
  STREET:  1185 Avenue of the Americas
  CITY:  New York
  STATE:  New York
  COUNTRY:  U.S.A.
  ZIP:  10036
COMPUTER READABLE FORM:
  MEDIUM TYPE:  Floppy disk
  COMPUTER:  IBM PC compatible
  OPERATING SYSTEM:  PC-DOS/MS-DOS
  SOFTWARE:  PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:  US/08/770,379
  FILING DATE:
  CLASSIFICATION:  435
ATTORNEY/AGENT INFORMATION:
  NAME:  White, John P.
  REGISTRATION NUMBER:  28,678
  REFERENCE/DOCKET NUMBER:  52342
TELECOMMUNICATION INFORMATION:
  TELEPHONE:  (212) 278-0400
  TELEFAX:  (212) 391-0525
INFORMATION FOR SEQ ID NO:  20:
  SEQUENCE CHARACTERISTICS:
  LENGTH:  32207 base pairs
  TYPE:  nucleic acid
  STRANDEDNESS:  double
  TOPOLOGY:  linear
  MOLECULE TYPE:  DNA (genomic)
US-08-770-379-20

Query Match        1.9%;   Score 60.2;  DB 2;  Length 32207;
Best Local Similarity   61.6%;   Pred. No. 0.00026;
Matches 114;  Conservative    0;  Mismatches  68;  Indels   3;  Gaps    1;

Qy   2998  CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTGCTAAACCT 3057
                 |||||  ||||||||||||||||||||||||||||||||||||||||| |||||
Db  19761  CATCCTGCTGCTGCTGCTCATCCTGCTCATCCTGCTCACCGCTCATCCT 19820

Qy   3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCGCTAAACCAACTGGGC-- 3115
                 ||||||| |||||||  ||||||||||||||||||||||||||||||||||||||
Db  19821  GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTCATCCTGCTGCTGCTGCTCCT 19880

Qy   3116  -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAAC 3174
                 |||||||| |||  |||||||||||||| |||||||||||||||| |||||
Db  19881  GCTCATCCTGCTCCTGCTCCTCGTGCTCCTGCTCATCCTGCTGCTCATCCTGCT 19940

Qy   3175  GCTCCC 3179
                 |||||
Db  19941  GCTGC 19945

RESULT 7
US-08-757-669A-20
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INFORMATION:
;  APPLICANT:  Chang, Yuan
;  APPLICANT:  Bohenzky, Roy A.
;  APPLICANT:  Russo, James J.
;  APPLICANT:  Edelman, Isidore S.
;  APPLICANT:  Moore, Patrick S.
;  TITLE OF INVENTION:  UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
;  TITLE OF INVENTION:  SEQUENCES AND USES THEREOF
;  NUMBER OF SEQUENCES:  20
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  Cooper & Dunham LLP
;  STREET:  1185 Avenue of the Americas
;  CITY:  New York
;  STATE:  New York
```

; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,669A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-757-669A-20

Query Match   1.9%;   Score 60.2; DB 3; Length 32207;
Best Local Similarity   61.6%;   Pred. No. 0.00026;
Matches 114; Conservative   0; Mismatches 68; Indels   3; Gaps   1;

Qy   2998 CCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCAAACCT 3057
            |||||||||||| ||||| |||||||||||||||||||||||||||||||||||||
Db   19761 CATCCTGCTGCTGCTCATCCTGCTCATCCTGCTGCTGCTGCTGCTGCTTCATCCT 19820

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   19821 GCTGCTGCTGCTCATCCTGCTGCTGCTCCTGCTCCTGCTCCTGCTGCTGCTCCT 19880

Qy   3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
             |||||||||||| ||| ||||| |||||||||||||||||||||||||||||
Db   19881 GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTGCTGCTGCTCATCCTGCT 19940

Qy   3175 GCTCC 3179
            |||||
Db   19941 GCTGC 19945

RESULT 8
US-09-230-371A-20
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match   1.9%;   Score 60.2; DB 4; Length 32207;
Best Local Similarity   61.6%;   Pred. No. 0.00026;
Matches 114; Conservative   0; Mismatches 68; Indels   3; Gaps   1;

Qy   2998 CCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCAAACCT 3057
            |||||||||||| ||||| |||||||||||||||||||||||||||||||||||||
Db   19761 CATCCTGCTGCTGCTCATCCTGCTCATCCTGCTGCTGCTGCTGCTGCTTCATCCT 19820

Qy   3058 GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC-- 3115
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   19821 GCTGCTGCTGCTCATCCTGCTGCTGCTCCTGCTCCTGCTCCTGCTGCTGCTCCT 19880

Qy   3116 -CTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACC 3174
             |||||||||||| ||| ||||| |||||||||||||||||||||||||||||
Db   19881 GCTCATCCTGCTGCTCCTGCTCATCCTGCTGCTGCTGCTGCTCATCCTGCT 19940

Qy   3175 GCTCC 3179
            |||||
Db   19941 GCTGC 19945

RESULT 9
US-09-790-988-1
; Sequence 1, Application US/0790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match   1.8%;   Score 58.4; DB 4; Length 640681;
Best Local Similarity   47.3%;   Pred. No. 0.0026;
Matches 209; Conservative   0; Mismatches 231; Indels   2; Gaps   1;

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTAAGT 300
            || | |||||||||| | ||| | || ||| | | ||| || |||| |||
Db   525047 ATTAAAAAAAAATAAAACACAGAAAATTTAATACATACATTTCTAAAAGTAATTTTATT 525106

Qy   301 GATGATTGTGGTGGATACAAAGTAAAACAAAATAGTTTCAGATTACACAACTAGCAGAAAT 360
            || |||| || | ||| | ||||| ||||| ||| ||| ||| | | || | ||||
Db   525107 TTAAAAAAGAAGAAGAAAAATATTATAAAACTATTATCTCAAAAAGAATATAAAAAA 525166

Qy   361 AGATTTGATCAAAGACAAACAAGAGACATATATGCTCTGTGTTGGTTAATGATGAAGCTAAC 420
            || ||| |||||| || ||||| | |||| || | | || ||||| || ||||| | ||
Db   525167 GCAATTGTTAAAATAGATCCTTCAATCTATCTAATAACAAATCAAGAAGCACAA 525226

Qy   421 GTTCATTTAAAAGAATAATAACTCAAATAGTTGGTAATAGAAAACAACAATTCT 480
            ||| || |||||| | |||||||||| | |||| | || |||| | |||||
Db   525227 ATTTATTTCAAAAAAAATCAAGATAATAATTTTTATATTTCCAGAAAAATTTAAAATTAATTTT 525286

Qy   481 AAGTTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAACTTACTGATGATGGG 540
            ||| ||| |||| | |||| ||| || ||| |||||||||| || || |||
Db   525287 GTTGAATTAAAACTGATAACTTTAAAATACATTGCGAAAATAAAGAGATTTATGATTGG 525346

Qy   541 ACTAAATTTAATTTACAAACCAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT 600
            ||| ||| || ||| | || ||| | | || ||| || | | || || || |||
Db   525347 TATATAGAAATATTAC--ACAATACTCAACTAAGAAAAAAGAGATATACTTATACATTCAAGAAAG 525404

Qy   601 GCGCAATCTTACCTAAAGATTTGGTATAACTTATACATTCAAGAAAG 660

Db   525405 AAGTCAAAAAATAAACAACAGCCATATCCGATATTTATCTAGATTGCATAATACCAGAAG 525464

Qy   661 ATCTTACCAAATGACGTCAACA 682

Db   525465 ATTTTTCAAAAAATAGCTCAAGA 525486

RESULT 10
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639053
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match   1.7%;   Score 54.2; DB 4; Length 832;
Best Local Similarity 13.1%;   Pred. No. 0.0015; Indels 0; Gaps 0;
Matches   47; Conservative 162; Mismatches 150; Indels 0; Gaps 0;

Qy   1348 AATCCTACTTTAAATACTTATCGTAGTTTTGGAATTGATGAGTAAACCAACTTCTGCAAAC 1407
                ||| |||| ||| || |||| : ::: || :
Db   390 AATTGTACCATTAAATAATTATATTTTGTYWTWNKTWYTTYTWTRMWWMKKKCARRWYY 331

Qy   1408 AAAATAGAGATGAAACTAATTGGGCAGATCCTAACGTTATTGAAGCAAGAATAATGCTGAA 1467
               :: || :   ::: :||| ::|::: ::| ::| : |:|::: |
Db   330 WWKSTYACASRYRKYTWGWWVWYWMKRMMSTRWYCYMCWKCCMVRGRRCAWYTWARGRMWS 271

Qy   1468 TACAGATTAGTATTCAAAATGAAAATTCCAATAACTAATGCAGGAAACTTTATCCGAAAC 1527
               :   :| :: ::| : : :: :| ::: ||:: ::| :
Db   270 YAWGKWKSMRSAMSMCTRMYYKKGSTYWTMKCTCATWCYWYWKYWKRMWSKTCWSGSRGG 211

Qy   1528 ACAATTGGTGCGTGTTGGTTTTACTTCAAGACAGTTCAAGAGAGTAGTTTTAAGAGCTTCTTAT 1587
               ::: :|||||||| ||| || | | | : |:::::
Db   210 YMTSYTSTRSYSMYWASWMYTMCWWWGRWNSTYWTMAWGKKWWRYATTWRRAMWWWAAW 151

Qy   1588 AACGGTGATCAACGTCCAACTGGAAACTTCCAACCTTTCTTATACGTATTTGGTTATTTA 1647
               | : ::   :|| ::: :: | :| ::|
Db   150 TMMWYMWWAWCMSSRGAAMYRRTMWWGYRYWWRKKSYRRTRCAWAYAWKTKRSYYWCWR 91

Qy   1648 GGATACCAACAAACTAGAACAGGAACTTCTGGTACGGAACATATAAGCTTTTAAACAA 1706
               :::: :|:::::|: :: |: : : | ::
Db   90 WKWKRCWMMMMAMAYGKTMMRACWKTRYWRWAWAMWRMWTWMMMYWYWWRAMKRR 32

RESULT 11
US-09-142-584-5
; Sequence 5, Application US/09142584
; Patent No. 6403094
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Havard, Helen L.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Payne, Dean W.
; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; FILE REFERENCE: 124-665
; CURRENT APPLICATION NUMBER: US/09/142,584

; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/GB97/00660
; EARLIER FILING DATE: 1997-03-11
; EARLIER APPLICATION NUMBER: GB 9605222.0
; EARLIER FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Clostridium perfringens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (136)..(987)
; FEATURE:
; NAME/KEY: misc_signal
; LOCATION: (1)..(32)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(984)
; FEATURE:
; OTHER INFORMATION: "n" at positions 451-453 represent a, t, c, g or
; OTHER INFORMATION: other
US-09-142-584-5

Query Match   1.7%;   Score 53.2; DB 4; Length 987;
Best Local Similarity 44.3%;   Pred. No. 0.0028; Indels 0; Gaps 0;
Matches 208; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

Qy   363 ATTTGATCAAGACAAACAAGAGCATATTATGCTCGTTGGTTAATGATGAGCTAACGT 422
               || | ||| | |:| | ||| || ||| :|| ||| ||| || ||| |:
Db   117 ATCTAATGAAATGTCCAAAAAAGCTTCTTATGATAATGTAGATACATCATTGAAGAAAGG 176

Qy   423 TCATTTAAAAAGAATTAATACTAACTCAAATGACTCAAATAGAAACAACAATTCTAA 482
               |||| |||| ||| | || ||| || | | || |||| ||| |||| ||
Db   177 AAGATATAATACAAAATAATAATTACTTAAAGAATGAAAAATATTATCCTAATGCTAT 236

Qy   483 GTTTGTAATTGGTGTGTTGATAATCCAGCTCACGTAATCAGTGGTTACTGATGATGGGAC 542
               ||| ||| || |||||| | | :|| :| | || | | ||||| |||| ||| ||
Db   237 GGCATATTTTGATAAGGTTACTATAAAATCCACAAGGAAATGAATTGATTTATTAATAATCC 296

Qy   543 TAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGC 602
               |||||| |||| ||| |||||||| ||||||||| |||| ||| || ||| | |||
Db   297 TAAAGTTGAATTAGAATGGAGGAACCATCAATGAATTATCTTGAAGATGTTTATGTTGGAAA 356

Qy   603 GCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGAT 662
               |||| | |||| ||| ||| ||||||||||| ||| || || |||| ||||||| |
Db   357 AGCTCTCTTAACTAAACTAATACTCAACAAGAACAAAAATTAAAATCACAATCATTCACTTG 416

Qy   663 CTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTTAGTGGAAC 722
               ||| | ||| || | | ||:| || || || |||| ||||| ||| ||| ||| |||
Db   417 TAAAAATACTGATACAGTACCTGCAACTACTCNNNACTGTGGGAACTTGATACAGC 476

Qy   723 AAATGCTGATGATGGGAATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTAT 782
               |||| |||||| || |: ||| |||| :|| |||| |||| ||| ||| ||| |||| |
Db   477 AACTGCTAAGTTACTGTTCCTTTTAATGAAACAGAATAATCAAATCCTTCAACTTTAAT 536

Qy   783 TGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTAAT 831
               ||| ||| | | ||| ||| ||| ||| |||| || ||||| ||| |
Db   537 TTTTGCAAATACAAATACAAATACTAATTCAAAAGAAATTACTCCATAAT 585

RESULT 12
US-09-181-585-3
; Sequence 3, Application US/09181585
; Patent No. 6524791
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 110090 0101
; CURRENT APPLICATION NUMBER: US/09/181,585
; CURRENT FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
; OTHER INFORMATION: comprising exons E, C, and A
US-09-181-585-3

Query Match   1.7%;   Score 53.2; DB 4;   Length 1037;
Best Local Similarity   50.8%;   Pred. No. 0.0028;
Matches 127; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTA 2989
            ||||| | ||| || ||
Db   605 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA 664

Qy   2990 AGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTG 3049
            |||||||||||||||||
Db   665 CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
            || ||||||||||||||||||||||||
Db   725 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784

Qy   3110 CTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAC 3169
            |||||||
Db   785 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844

Qy   3170 CAACCGCTCC 3179
            ||||||||||
Db   845 CTGCTGCTGC 854

RESULT 13
US-09-181-585-1
; Sequence 1, Application US/09181585
; Patent No. 6524791
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/09/181,585
; CURRENT FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-181-585-1

Query Match   1.7%;   Score 53.2; DB 4;   Length 1159;
Best Local Similarity   50.8%;   Pred. No. 0.003;
Matches 127; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTA 2989
            ||||| | ||| || ||
Db   417 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA 476

Qy   2990 AGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTG 3049
            |||||||||||||||||
Db   477 CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
            || ||||||||||||
Db   537 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596

Qy   3110 CTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAC 3169
            ||||||| |||||||
Db   597 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656

Qy   3170 CAACCGCTCC 3179
            ||||| |||
Db   657 CTGCTGCTGC 666

RESULT 14
US-09-181-585-2
; Sequence 2, Application US/09181585
; Patent No. 6524791
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/09/181,585
; CURRENT FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1471
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
; OTHER INFORMATION: comprising exons D, C, B, and A
US-09-181-585-2

Query Match   1.7%;   Score 53.2; DB 4;   Length 1471;
Best Local Similarity   50.8%;   Pred. No. 0.0033;
Matches 127; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   2930 CCTTGACTGCTGCTGTTGGTTCAGTTACAAGAAGATTATTACCCAAACTGCTAACGTTA 2989
            ||||| | ||| || ||
Db   1039 CCTGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTACTA 1098

Qy   2990 AGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTG 3049
            |||||||||||||||||
Db   1099 CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158

Qy   3050 CTAAACCTGCTGCTCCAGCTAAACCATCTGACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
            || ||||||||||||
Db   1159 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218

Qy   3110 CTGGGGCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAAC 3169
            |||||||
Db   1219 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278

Qy   3170 CAACCGCTCC 3179
            ||||||
Db   1279 CTGCTGCTGC 1288

RESULT 15
US-09-142-584-3
; Sequence 3, Application US/09142584
; Patent No. 6403094
; GENERAL INFORMATION:
; APPLICANT: Titball, Richard W.
; APPLICANT: Williamson, Ethel D.
; APPLICANT: Havard, Helen L.
; APPLICANT: Oyston, Petra C.F.
; APPLICANT: Payne, Dean W.
; TITLE OF INVENTION: CLOSTRIDIUM PERFRINGENS VACCINES
; FILE REFERENCE: 124-665
; CURRENT APPLICATION NUMBER: US/09/142,584
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/GB97/00660
; EARLIER FILING DATE: 1997-03-11
; EARLIER APPLICATION NUMBER: GB 9605222.0
; EARLIER FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 6

```
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 3
;  LENGTH: 987
;  TYPE: DNA
;  ORGANISM: Clostridium perfringens
;  FEATURE:
;  NAME/KEY: mat_peptide
;  LOCATION: (136)..(987)
;  FEATURE:
;  NAME/KEY: misc_signal
;  LOCATION: (1)..(32)
;  FEATURE:
;  NAME/KEY: CDS
;  LOCATION: (1)..(984)
US-09-142-584-3
```

```
Query Match         1.7%;   Score 53;  DB 4;  Length 987;
Best Local Similarity  44.6%;  Pred. No. 0.0031;
Matches 209;  Conservative  0;  Mismatches 260;  Indels  0;  Gaps  0;

Qy    363  ATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGT  422
                |  | || | |||    |||  |    ||||||| |   | | |||||| | ||
Db    117  ATCTAATGAAAATGTCCAAAAAAAGCTTCTTATGATAAATGTAGATACATTAATTGAGAGAAGG  176

Qy    423  TCATTTAAAAAGAATTAAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAA  482
                | ||| | | |||  | | |||| || ||||  | || | |||| | || ||| |
Db    177  AAGATATAATACAAAATATTAATTACTTAAAGAGAATGGAAAAAATATTATCCTAATGCTAT  236

Qy    483  GTTTGTAAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGAC  542
                | |  | | |||| ||| || |  | | | ||||  | | || |||| |||||
Db    237  GGCATATTTTGATAAGGTTACTATAAAATCCACAAGGAAATGATTTTTATATTAATAATCC  296

Qy    543  TAAATTTAATTTTACAAACCAAACTCAAGGTGAAAATTGTTAATGACTTCATTTTAGATGC  602
                ||||   |||||   |    ||  || ||||| |  || ||  ||| ||
Db    297  TAAAGTTGAATTAGATGGGAGAACCATCAATGAATTATCTTGAGATGTTTATGTTGGAAA  356

Qy    603  GCCAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTAGCATTCAAAGAAAGAT  662
                |  ||  | ||| |   | |   | || ||  | ||  |  |||| ||| |
Db    357  AGCTCTCTTAACTAATGATACTCAACAAGAACAAAAATTAAAATCACAATCATTCACTTG  416

Qy    663  CTTACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTTAGTGGAAC  722
                || |    |  | | || |    ||  |  | || ||||| ||| | |||||
Db    417  TAAAAATACTGATACAGTAACTGCAACTACTACTCCGACTGTGGGAACTTCGATACAAGC  476

Qy    723  AAATGCTGATGATGGGATGTTTGATTGTGGGGAATGGTCAAATAACTAATACAGAATCCTAT  782
                |||| | | | | ||| ||  | || |  | | | | || | | || |||||||
Db    477  AACTGCTAAGTTTACTGTCTTCCTTTTAATGAAACAGGAGTATCATTAACTACTAGTTATAG  536

Qy    783  TGCTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTAAT  831
                ||| ||| | | || |  | |   ||| | | ||||
Db    537  TTTTGCAAATACAAATACAAATACTAATTCAAAGAAATTACTCATAAT  585
```

Search completed: May 5, 2004, 20:04:15
Job time : 247.505 secs

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2004   Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May  5, 2004, 17:07:04 ; Search time 1255.74 Seconds
                                    (without alignments)
                                    11489.234 Million cell updates/sec

Title:          US-09-901-572A-2
Perfect score:  3189
Sequence:       1 atgaatatttctaaaaact.........caaccgctccaaagaataa 3189

Scoring table:  IDENTITY NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2937390 seqs, 226206796 residues

Total number of hits satisfying chosen parameters:   5874780

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published Applications NA:*
   1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
   2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
   3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
   4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
   5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
   6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
   7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
   8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
   9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
  10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
  11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
  12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
  13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
  14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
  15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
  16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
  17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
  18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
  19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	3189	100.0	3189	10	US-09-901-572A-2	Sequence 2, Appli
2	3189	100.0	3189	15	US-10-131-591A-2	Sequence 2, Appli
3	3143.4	98.6	8354	14	US-10-125-818-1	Sequence 79, Appl
4	3057.8	95.9	3129	15	US-10-131-591A-79	Sequence 79, Appl
5	124.6	3.9	3159	13	US-10-282-122A-27316	Sequence 27316, A
6	124.6	3.9	580073	15	US-10-205-220-1	Sequence 1, Appli
7	62.6	2.0	4985	15	US-10-094-240-10	Sequence 10, Appl
8	62.6	2.0	4985	15	US-10-056-405-10	Sequence 10, Appl
9	60.8	1.9	8771	13	US-10-311-455-1798	Sequence 1798, Ap
10	60.2	1.9	3489	15	US-10-894-273-1	Sequence 1, Appli
11	60.2	1.9	3489	9	US-10-294-804-1	Sequence 1, Appli
12	58.4	1.8	640681	9	US-09-790-988-1	Sequence 1, Appli
13	53.6	1.7	3673778	15	US-10-312-841-1	Sequence 1, Appli
14	53.2	1.7	1037	16	US-10-373-667-3	Sequence 3, Appli
15	53.2	1.7	1159	16	US-10-373-667-1	Sequence 1, Appli
16	53.2	1.7	1471	16	US-10-373-667-2	Sequence 2, Appli
17	53	1.7	13784	17	US-10-257-166-144	Sequence 144, App
18	52.4	1.6	3057	15	US-10-349-680-148	Sequence 148, App
19	51.6	1.6	906	13	US-10-282-122A-16807	Sequence 16807, A
20	51.6	1.6	2451	15	US-10-091-007-55	Sequence 55, Appl
21	51.4	1.6	2553	13	US-10-282-122A-40676	Sequence 40676, A
22	50.8	1.6	2017	15	US-10-155-533-3	Sequence 3, Appli
23	50.8	1.6	5314	15	US-10-155-533-1	Sequence 1, Appli
24	50.2	1.6	2010	9	US-09-878-756-3	Sequence 3, Appli
25	49.8	1.6	486	13	US-10-465-217-13	Sequence 13, Appl
26	49.4	1.5	1887	15	US-10-349-680-133	Sequence 133, App
27	49	1.5	718	16	US-10-264-049-422	Sequence 422, App
28	48.6	1.5	357	13	US-10-085-783A-7843	Sequence 7843, Ap
29	48.6	1.5	357	16	US-10-242-535A-7843	Sequence 7843, Ap
30	48	1.5	720	13	US-10-424-599-129163	Sequence 129163,
31	48	1.5	2183	16	US-10-104-047-1064	Sequence 1064, Ap
32	47.8	1.5	3931	15	US-10-006-780-1	Sequence 1, Appli
33	47.6	1.5	1220	15	US-10-169-710-46	Sequence 46, Appl
34	47.4	1.5	972	13	US-10-243-552-209	Sequence 209, App
35	47.2	1.5	769	13	US-10-363-616-47	Sequence 47, Appl
36	47	1.5	303	13	US-10-085-783A-33947	Sequence 33947, A
37	47	1.5	303	16	US-10-242-535A-33947	Sequence 33947, A
38	47	1.5	364	13	US-10-085-783A-15800	Sequence 15800, A
39	47	1.5	364	16	US-10-244-535A-15800	Sequence 15800, A
40	47	1.5	375	10	US-09-918-995-6127	Sequence 6127, Ap
41	47	1.5	712	9	US-09-822-849A-186	Sequence 186, App
42	47	1.5	722	9	US-09-923-304-1	Sequence 1, Appli
43	47	1.5	722	9	US-09-880-107-1764	Sequence 1764, Ap
44	47	1.5	755	13	US-10-027-632-128109	Sequence 128109,
45	47	1.5	755	16	US-10-027-632-128109	Sequence 128109,

ALIGNMENTS

```
RESULT 1
US-09-901-572A-2
; Sequence 2, Application US/09901572A
; Publication No. US20030165534A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/09/901,572A
; CURRENT FILING DATE: 2003-03-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3189
; TYPE: DNA
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: mgc3 gene
US-09-901-572A-2

Query Match         100.0%;  Score 3189;  DB 10;  Length 3189;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 3189;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGATTAGCTGTATTGGA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAATATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGATTAGCTGTATTGGA 60

Qy         61 GCTCTTGGTTCTGCGAGCTTGGCTTTAAGCAATCAGATACAGATAAGCAACGATAACACGCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTCTTGGTTCTGCGAGCTTGGCTTTAAGCAATCAGATACAGATAAGCAACGATAACACGCAA 120

Qy        121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTAGACTTGGACTTCGCAGGTCTTGGACA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCGTTAGACTTGGACTTCGCAGGTCTTGGACA 180

Qy        181 AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAAGCAGCAGCT 240
```

Left column:

```
Db   181 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         AATGGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCT  240

Qy   241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 AATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGGTTTAGATCTAAGT  300

Qy   301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCACAACTAGCAGAAAT  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GATGATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACCACAACTAGCAGAAAT  360

Qy   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGATTTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAAC  420

Qy   421 GTTCATTAAAAAGAATTAATAACTCAAATAGAATTGGTAATAGAACAACAATTCT  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GTTCATTAAAAAGAATTAATAACTCAAATAGAATTGGTAATAGAACAACAATTCT  480

Qy   481 AAGTTTGTAATTGGTGGTGTACAAGTAAACCAGCTCACGTAATTACTGATGATGGG  540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 AAGTTTGTAATTGGTGGTGTACAAGTAAACCAGCTCACGTAATTACTGATGATGGG  540

Qy   541 ACTAAATTAATTTTACAAACCAAGTGAAATTGTAATGACTTCATTTTAGAT  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 ACTAAATTAATTTTACAAACCAAGTGAAATTGTAATGACTTCATTTTAGAT  600

Qy   601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660

Qy   661 ATCTTACCAAATGACGTCAACACTGACTTGTTCCTTGGCCAGTAGGTAGTGA  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATCTTACCAAATGACGTCAACACTGACTTGTTCCTTGGCCAGTAGGTAGTGA  720

Qy   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 ACAAATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCT  780

Qy   781 ATTGCTCAAACTAAAACCACTACTGATAATAATCAAAATCCTTCAACTTTAATTCAGGAGCA  840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATTGCTCAAACTAAAACCACTACTGATAATAATCAAAATCCTTCAACTTTAATTCAGGAGCA  840

Qy   841 ATGCCTGGTGCAAAAATAGATACGATTCTCAAGTCAAGCATAGAAATTAAAACA  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 ATGCCTGGTGCAAAAATAGATACGATTCTCAAGTCAAGCATAGAAATTAAAACA  900

Qy   901 TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TCTTTCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAA  960

Qy   961 AATATTACAAGATTAGCTACGGGAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 AATATTACAAGATTAGCTACGGGAGTTGCCAAGCAACGAAAGATATTGGATTCTTGAC  1020

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGAATTCAGTTAACGTAACGTAACAACAA  1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGAATTCAGTTAACGTAACGTAACAACAA  1080

Qy  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTAGATAGTATTTATATTTTGGTACC  1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TACTTAAACTCAGTTAATTCTTTATCATTCATTGGTAGATAGTATTTATATTTTGGTACC  1140

Qy  1141 TCTGAATTACAAGTTAAATCACATTATGGTACTATTCCCAACTAGATTATCTGATCTAACCGCT  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TCTGAATTACAAGTTAAATCACATTATGGTACTATTCCCAACTAGATTATCTGATCTAACCGCT  1200

Qy  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACACAACAACA  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 TTGAATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACACAACAACA  1260

Qy  1261 AACGGAACACGACAACGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAGAAAT  1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column:

```
Db  1261 AACGGAACACGACAACGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAGAAAT  1320

Qy  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGA  1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACTACTAACACTTCTCAAACAGTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGA  1380

Qy  1381 ATTGATAGTAAACCAACTTCTGCAACAACAAATAGATGAAACTAATTGGGCAGATCCTAAC  1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 ATTGATAGTAAACCAACTTCTGCAACAAAAATAGATGAAACTAATTGGGCAGATCCTAAC  1440

Qy  1441 GTTATTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA  1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 GTTATTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATA  1500

Qy  1501 ACTAATGCAGGAAACTTTATCCGAAACACAATGGTGGTGTTGTTTACTTCAACAGGT  1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ACTAATGCAGGAAACTTTATCCGAAACACAATGGTGGTGTTGTTTACTTCAACAGGT  1560

Qy  1561 TCAAGAGTAGTTTAAGAGCTTCTATAAACGTGATCAACTGGAAACTTCCAA  1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TCAAGAGTAGTTTAAGAGCTTCTATAAACGTGATCAACTGCCAACTTCCAA  1620

Qy  1621 CCTTTCTTATACGTATTTGGTTATTAGGATACCAACAAACTAGAACAGGAACTTTCTGG  1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CCTTTCTTATACGTATTTGGTTATTAGGATACCAACAAACTAGAACAGGAACTTTCTCG  1680

Qy  1681 TACGGAACATATAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTCAAGAGTA  1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACGGAACATATAGCTTTAAACAACAGCCCTTACGACGTATTAGATTCTCAAGAGTAT  1740

Qy  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACAATTAAGAAGCACAAGGTGAC  1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 GGTACTGAAACCAATCAATTAGAAGAACTTCATTAACAATTAAGAAGCACAAGGTGAC  1800

Qy  1801 CTAACTGAAGAAGGTGCAGAGAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT  1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 CTAACTGAAGAAGGTGCAGAGAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT  1860

Qy  1861 ACACCGAGAAAGCCGAAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT  1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACACCGAGAAAGCCGAAGCATCTTCCAATCTGCTATTCTGATAATACTTATGAGTACATT  1920

Qy  1921 CAATCAGTTTTAGGATTTGATGAATTAGAATTAGAAATTAGAAGAATAACTTAAAAGTGTTGGGGTTAAAGCATCA  1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CAATCAGTTTTAGGATTTGATGAATTAGAATTAGAAATTAGAAGAATAACTTAAAAGTGTTGGGGTTAAAGCATCA  1980

Qy  1981 AGTTCTTCAAATCAAATAGACCAAATCCAAACGTCTAGAAATGATTGCTGCAACAACA  2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 AGTTCTTCAAATCAAATAGACCAAATCCAAACGTCTAGAAATGATTGCTGCAACAACA  2040

Qy  2041 TACTTAAGATCACAAATTGGATTAGCTGACATCTGGATAGATTACCAACAACAACCATTC  2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 TACTTAAGATCACAAATTGGATTAGCTGACATCTGGATAGATTACCAACAACAACCATTC  2100

Qy  2101 GGAACAACTCACCAGTTATTTCAGTATCAGCCTGGTGATCAGTCTCATCAATTAAGAAT  2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 GGAACAACTCACCAGTTATTTCAGTATCAGCCTGGTGATCAGTCTCATCAATTAAGAAT  2160

Qy  2161 ATTGAAACAATCTTCCCTGGTAACCAGTTAATGGTGACTTCTTATTCAACAAATGAAAATAAT  2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTAATGGTGACTTCTTATTCAACAAATGAAAATAAT  2220

Qy  2221 AAAATCTAGTGTTTATACATTAAGATTAGATAGCTGACTCAACGTAACCGATGCGTCAAGCTCA  2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AAATCTAGTGTTTATACATTAAGATTAGATAGCTGACTCAAGTGACCGATGCGTCAAGCTCA  2280

Qy  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGAATCTTTACCTTTATTAGAC  2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGAATCTTTACCTTTATTAGAC  2340

Qy  2341 AATTCATTCATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGT  2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 AATTCATTCATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGT  2400
```

; ORGANISM: Mycoplasma galliepticum
; FEATURE:
; OTHER INFORMATION: mgc3 gene
US-10-131-591A-2

Query Match   100.0%;   Score 3189;   DB 15;   Length 3189;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 3189;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1 ATGAATATTTCTAAAAAACTAAAAGTATACATTGATAGTGGATTAGCTGTATTTGGA   60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAATATTTCTAAAAAACTAAAAGTATACATTGATAGTGGATTAGCTGTATTTGGA   60

Qy   61 GCTCTTGGTTCTGCAAGCTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 GCTCTTGGTTCTGCAAGCTTGGCTTAAGCAATCAGATAAGAGTAACGATAACACGCAA  120

Qy  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAA  180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTAGTTAATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAA  180

Qy  181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGAATAACTTTATAACAGCAGCT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AATGGTTCGTGTTCAATACAGTTCTTAGAGATGTTGATGAATAACTTTATAACAGCAGCT  240

Qy  241 AATGGAACAATTATCAAATTAGATAGTTTTACTCAGCAATTGGTAGTACATTA  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 AATGGAACAATTATCAAATTAGATAGTTTTACTCAGCAATTGGTAGTACATTA  300

Qy  301 GATGATTGGTGGATCACAAGAGTAAAACAAAATAGTTCAGATTACACAACTAGCACAGAAAT  360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GATGATTGGTGGATCACAAGAGTAAAACAAAATAGTTCAGATTACACAACTAGCACAGAAAT  360

Qy  361 AGATTTGATCAAAGACAAACAAGACAAGAGCATATTATGCTCTCTGTTGGTTAATGAAGCTAAC  420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AGATTTGATCAAAGACAAACAAGACAAGAGCATATTATGCTCTCTGTTGGTTAATGAAGCTAAC  420

Qy  421 GTTCATTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCT  480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GTTCATTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCT  480

Qy  481 AAGTTTGTAATTGGTTGGTGTGTGATAATCCAGCTCACGTAATTAGAGATTTACTGATGATGGG  540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 AAGTTTGTAATTGGTTGGTGTGTGATAATCCAGCTCACGTAATTAGAGATTTACTGATGATGGG  540

Qy  541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGAT  600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGAT  600

Qy  601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GCGCCAATCTTACCTAAAGATTACACCCAGATTGGTATAACTTATACATTCAAAGAAAG  660

Qy  661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTGATTGTGGAATGGTCAAATAACTAATAACATAGATCCT  720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ATCTTACCAAATGACGTCAACACTGCAGTTGTTGATTGTGGAATGGTCAAATAACTAATAACATAGATCCT  720

Qy  721 ACAAATGCTGATGATGGGAATGTTGATTGTGGGGAATGGTCAAATAACTAATAACATACAGGAGCA  780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 ACAAATGCTGATGATGGGAATGTTGATTGTGGGGAATGGTCAAATAACTAATAACATACAGGAGCA  780

Qy  781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAAATCCTCAACTTTTAATTCAGGAGGACA  840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 ATTGCTCAAACTAAAACCACTACTGATAATCAAAAATCCTCAACTTTTAATTCAGGAGGACA  840

Qy  841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA  900
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 ATGCCTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACA  900

Qy  901 TCTTTCCAATTAGATGATGAAAAATTTGTTTATCCGAGAATGGACTGGTTCTGAAGAGAATAA  960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TCTTTCCAATTAGATGATGAAAAATTTGTTTATCCGAGAATGGACTGGTTCTGAAGAGAATAA  960
```

```
Qy 2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTATCTGATATTGCTTTT 2460

Qy 2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 GAAGGTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGAT 2520

Qy 2521 GAGTACTTAATTCAAATGGGTTCACTAGTCAGTGGCTAGAAACTTCGTTACAAACCAA 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 GAGTACTTAATTCAAATGGGTTCACTAGTCAGTGGCTAGAAACTTCGTTACAAACCAA 2580

Qy 2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 AGCTTCTTAAACAGTTTAGTTGACTTCACTCCTGCTAATGCTGGTACTAACTACCGTGTA 2640

Qy 2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTTACCTCTAAAAGTTCAGATCCAA 2700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 GTGGTTGATCCTGATGGTAATTTAACAAACCAAAACTTACCTCTAAAAGTTCAGATCCAA 2700

Qy 2701 TACTTAGATGTGTAAGTATTATGATGTCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 TACTTAGATGTGTAAGTATTATGATGTCTAAATTAAAGAACAATAATTTAGTAACATTCTCT 2760

Qy 2761 TATAACAACTTGGGGCCTTACACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 TATAACAACTTGGGGCCTTACACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTA 2820

Qy 2821 GGTATTCTTCTGCAATTATGATCATCTTTAGGATTAGGTATCGGTATCGGTTTTAAGAGCTCAA 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 GGTATTCTTCTGCAATTATGATCATCTTTAGGATTAGGTATCGGTATCGGTTTTAAGAGCTCAA 2880

Qy 2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT 2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 AGAAAATTACAAGACAAAGGGTTCAAAACAACATTCAAAAAGTTGATACCTTGACTGCT 2940

Qy 2941 GCTGTGGTTCAGTTTACAGAAGATTATTACCCAACTGCTACGTTAAGAGAAAAAACCT 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 GCTGTGGTTCAGTTTACAGAAGATTATTACCCAACTGCTACGTTAAGAGAAAAAACCT 3000

Qy 3001 GCTGCTTAGGTGCTGGTAAATCTGTGGTAAGAACCTGCTCGCTGCTAAACCTGCT 3060
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 GCTGCTTAGGTGCTGGTAAATCTGTGGTAAGAACCTGCTCGCTGCTAAACCTGCT 3060

Qy 3061 GCTCCAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 GCTCCAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAAACCAACTGGGCCTAAA 3120

Qy 3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACAACCGCTCCC 3180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 TCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACAACCGCTCCC 3180

Qy 3181 AAAGAATAA 3189
         |||||||||
Db 3181 AAAGAATAA 3189
```

RESULT 2
US-10-131-591A-2
; Sequence 2, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 3189
; TYPE: DNA

```
Qy  2041 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 2100
Db  2041 TACTTAAGATCACAAATTGGATTAGTAGCTAGAAACATCTGATTACCAAACCAACCACCATTC 2100

Qy  2101 GGACAACTCACCAAGTTATTCAGTATCAGTTCTCATCAGTTCTCATCAATTAAGAAT 2160
Db  2101 GGACAACTCACCAAGTTATTCAGTATCAGTTCTCATCAGTTCTCATCAATTAAGAAT 2160

Qy  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAT 2220
Db  2161 ATTAGAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAAT 2220

Qy  2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280
Db  2221 AAATCTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCA 2280

Qy  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTATTAGAC 2340
Db  2281 TTCAGTCCAACAAGTTAATTGACGTTAATGAAATTGGTGTAATCTTATTAGAC 2340

Qy  2341 AATTCATTCTATACAGTAAAATGCTGCTGGTAATCTGTTCTCATCAAACCCTGCT 2400
Db  2341 AATTCATTCTATACAGTAAAATGCTGCTGGTAATGTGTTCTCATCAAACCCTGT 2400

Qy  2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTAATCTGATATTGCTTTT 2460
Db  2401 TCTCCTGGATCATATACTGCTGTAAATACATTAATCAGAACTTAATCTGATATTGCTTTT 2460

Qy  2461 GAAGGTTCTGGTGCTGAAGTATACATCTGATTTCTGGGGAACAATCAAATCAAACCCGAT 2520
Db  2461 GAAGGTTCTGGTGCTGAAGTATACATCTGATTTCTGGGGAACAATCAAATCAAACCCGAT 2520

Qy  2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580
Db  2521 GAGTACTTAATTCAAAATGGGTTCACTAGTCAAGTGGCTAGAAACTTCGTTACAAACCAA 2580

Qy  2581 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTCGGTACTAACTACCGTGTA 2640
Db  2581 AGCTTCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTCGGTACTAACTACCGTGTA 2640

Qy  2641 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTCTAAAAGTTCAGATCCAA 2700
Db  2641 GTGGTTGATCCTGATGGTAATTAACAAACCAAAACCTCTAAAAGTTCAGATCCAA 2700

Qy  2701 TACTTAGATGGTAAGTAATATGATGTAATAAATAAAGAACAATAATTAGTAACATTCTCT 2760
Db  2701 TACTTAGATGGTAAGTAATATGATGTAATAAATAAAGAACAATAATTAGTAACATTCTCT 2760

Qy  2761 TATAACAACTTGGGCGCCTTCATGGGCCTTACCTTCATGGGCCTTACAGCAATTGGTAGTACATTA 2820
Db  2761 TATAACAACTTGGGCGCCTTCATGGGCCTTACCTTCATGGGCCTTACAGCAATTGGTAGTACATTA 2820

Qy  2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCTTCCTTAAGAGCTCAA 2880
Db  2821 GGTATTCTTGCAATTATGATCATCTTAGGATTAGCTCTTCCTTAAGAGCTCAA 2880

Qy  2881 AGAAAATTACAGACAAAGGGTTCAAAACAATTCAAAAAAGTTGATACCTTGACTGCT 2940
Db  2881 AGAAAATTACAGACAAAGGGTTCAAAACAATTCAAAAAAGTTGATACCTTGACTGCT 2940

Qy  2941 GCTGTTGGTTCAGTTTACAAGAAGATTAATTACCCAAAACTGCTAACGTTAAGAAAAAACCT 3000
Db  2941 GCTGTTGGTTCAGTTTACAAGAAGATTAATTACCCAAAACTGCTAACGTTAAGAAAAAACCT 3000

Qy  3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGTCTAAACCTGCTGCT 3060
Db  3001 GCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGTCTAAACCTGCT 3060

Qy  3061 GCTCCAGCTAAACCATCTGACCACCAAAAGCTAGTCGACCAGTAAACCAACTGGGCCTAAA 3120
Db  3061 GCTCCAGCTAAACCATCTGACCACCAAAAGCTAGTCGACCAGTAAACCAACTGGGCCTAAA 3120

Qy  3121 TCTGGTGCGCCTACAAAACCCTACAACCTGCTCTCAAGCAGCTGCTCCAAACCGCTCCC 3180
```

```
Qy   961 AATATTACAGATTAGCTACTGGGAGTTTGCCAAGCAACGAAGAGATATTGGATTCTTGAC 1020
Db   961 AATATTACAGATTAGCTACTGGGAGTTTGCCAAGCAACGAAGAGATATTGGATTCTTGAC 1020

Qy  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAAGCGTATTTTCAAGACTA 1080
Db  1021 ATACCCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAAGCGTATTTTCAAGACTA 1080

Qy  1081 TACTTAAACTTCAGTTAATTCTTTATCATTCATTCGTGATAGTATTTTATATTTTTGGTACC 1140
Db  1081 TACTTAAACTTCAGTTAATTCTTTATCATTCATTCGTGATAGTATTTTATATTTTTGGTACC 1140

Qy  1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAATCTGATCTAACCGGCT 1200
Db  1141 TCTGAATTACCATCATTATGGTACTATTCATTCCCAACTAATCTGATCTAACCGGCT 1200

Qy  1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260
Db  1201 TTGAATCAAGTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACA 1260

Qy  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320
Db  1261 AACGGAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAAT 1320

Qy  1321 ACTACTACACTTCTCAAACAGTTTCTAAATACTTAAATACTTATCGTAGTTTTGGA 1380
Db  1321 ACTACTAACACTTCTCAAACAGTTTCTAAATACTTAAATACTTATCGTAGTTTTGGA 1380

Qy  1381 ATTGATAGTAAAACCAACTTCTGCAAACAAATAGATGAAAATAATTGGGCAGATCCTAAC 1440
Db  1381 ATTGATAGTAAAACCAACTTCTGCAAACAAATAGATGAAAATAATTGGGCAGATCCTAAC 1440

Qy  1441 GTTATTGAAGCAAGAATAATATGCTGATACATCTAGTGGTATTAGTAGGTATTCAAATA 1500
Db  1441 GTTATTGAAGCAAGAATAATATGCTGATACATCTAGTGGTATTAGTAGGTATTCAAATA 1500

Qy  1501 ACTAATGCGAGGAAAACTTTATCCGAAACACAATTGGTGTGTTACTTCAACAGGT 1560
Db  1501 ACTAATGCGAGGAAAACTTTATCCGAAACACAATTGGTGTGTTACTTCAACAGGT 1560

Qy  1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGTGATCAACGTCCAACTGGAAACTTCCAA 1620
Db  1561 TCAAGAGTAGTTTTAAGAGCTTCTTATAACGTGATCAACGTCCAACTGGAAACTTCCAA 1620

Qy  1621 CCTTTCTTATACGTATTTGGTATTTAGGATCAACAACAAACAGAACAGGAACTTTCTGG 1680
Db  1621 CCTTTCTTATACGTATTTGGTATTTAGGATCAACAACAAACAGAACAGGAACTTTCTGG 1680

Qy  1681 TACGGAACATATAAAGCTTTAAACAACAGCCCTTACGACGTATTAGTTCTCCCAGAGTA 1740
Db  1681 TACGGAACATATAAAGCTTTAAACAACAGCCCTTACGACGTATTAGTTCTCCCAGAGTA 1740

Qy  1741 GGTACTGAAACCAATCAATTTAGGATTAGAATTAGAAATAACTTCATTAAATGTTGGGTTAAGCATCA 1800
Db  1741 GGTACTGAAACCAATCAATTTAGGATTAGAATTAGAAATAACTTCATTAAATGTTGGGTTAAGCATCA 1800

Qy  1801 CTAACTGAAGAGGTGCTAGAAGTTTCTAATATCCTAATACTCCATATATAAGAGCACAAGGTGAC 1860
Db  1801 CTAACTGAAGAGGTGCTAGAAGTTTCTAATATCCTAATACTCCATATATAAGAGCACAAGGTGAC 1860

Qy  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATCTTATGAGTACATT 1920
Db  1861 ACACCAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATCTTATGAGTACATT 1920

Qy  1921 CAATCAGTTTGATGATGAATTAGAAATAACTTAAAATGTTGGGGTTAAAGCATCA 1980
Db  1921 CAATCAGTTTGATGATGAATTAGAAATAACTTAAAATGTTGGGGTTAAAGCATCA 1980

Qy  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACCGTCACAGGTCGTCTAGAAATGATTGCTGCAACAACA 2040
Db  1981 AGCTTCTTAAACTCAAATAGACCAAATCCAAACCGTCACAGGTCGTCTAGAAATGATTGCTGCAACAACA 2040

Qy  2041 TACTTAAGATCACAAATTGGATTAGTAGCTAGAAACATCTGATTACCAAACCAACCACCATTC 2100
```

```
Db  3121 TCTGGTGCGCTACAAAACCAACTGCTCCTCAAGCCAGCTGCTCCAAAACCAACCGCTCCC 3180

Qy  3181 AAAGAATAA 3189
              ||||||||
Db  3181 AAAGAATAA 3189

RESULT 3
US-10-125-818-1
; Sequence 1, Application US/10125818
; Publication No. US20020187162A1
GENERAL INFORMATION:
;  APPLICANT: University of Connecticut
;  APPLICANT: Geary, Steven J.
;  APPLICANT: Sekellick, Margaret
;  APPLICANT: Marcus, Philip J.
;  APPLICANT: Silbart, Lawrence
;  TITLE OF INVENTION: Use of a Live Attenuated Mycoplasma gallisepticum Strain as a Vac
;  TITLE OF INVENTION: Vector for the Protection of Chickens and Turkeys from Respirato
;  FILE REFERENCE: 883933.0074
;  CURRENT APPLICATION NUMBER: US/10/125,818
;  CURRENT FILING DATE: 2002-04-19
;  PRIOR APPLICATION NUMBER: 60/285,569
;  PRIOR FILING DATE: 2001-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;  LENGTH: 8354
;  TYPE: DNA
;  ORGANISM: Mycoplasma gallisepticum
US-10-125-818-1

Query Match       98.6%;  Score 3143.4;  DB 14;  Length 8354;
Best Local Similarity   99.1%;  Pred. No. 0;
Matches 3159;  Conservative   2;  Mismatches   28;  Indels   0;  Gaps   0;
```

RESULT 4
US-10-131-591A-79
; Sequence 79, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.,
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131,591A
; CURRENT FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 3129
; TYPE: DNA
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: Modified mgc3 gene
US-10-131-591A-79

Query Match   95.9%; Score 3057.8; DB 15; Length 3129;
Best Local Similarity   98.7%; Pred. No. 0;
Matches 3083; Conservative   0; Mismatches 42; Indels   0; Gaps   0;

(DNA sequence alignment blocks, Query (Qy) vs Database (Db), with numbered positions on left and right margins)

```
Qy     305 ATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGAT 364
           ||| || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     245 ATTTTGCTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGAT 304

Qy     365 TTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTC 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 TTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTC 364

Qy     425 ATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGT 484
           ||||||||||||||||||||||||||||||||||||||||||||||| | ||||||||||
Db     365 ATTTAAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAAACCAAAATTCTAAGT 424

Qy     485 TTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTA 544
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     425 TTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTA 484

Qy     545 AATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGC 604
           ||||| | |||||| | |||||||||||||||||||||||||||||||||||||||||||
Db     485 AATTTCAGTTTACACAGCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGC 544

Qy     605 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCT 664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     545 CAATCTTACCTAAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAAGATCT 604

Qy     665 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGAACAA 724
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     605 TACCAAATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTAGTGGAACAA 664

Qy     725 ATGCTGATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTG 784
           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db     665 ATGCTGATGATGGGATGTTTGATTTTGGGAATGGTCAAATAACTAATACAGATCCTATTG 724

Qy     785 CTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC 844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     725 CTCAAACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGC 784

Qy     845 CTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     785 CTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT 844

Qy     905 TCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAATA 964
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db     845 TCCAATTAGATGAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAACAGA 904

Qy     965 TTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     905 TTACAAGATTAGCTACTGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGACATAC 964

Qy    1025 CCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTATACT 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     965 CCGGGACTCCACAAGTTACTTTAAAAGAAGATTCAGTTAACGTATTTTCAAGACTATACT 1024

Qy    1085 TAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTTATATTTTTGGTACCTCTG 1144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1025 TAAACTCAGTTAATTCTTTATCATTCATTGGTGATAGTATTTATATTTTTGGTACCTCTG 1084

Qy    1145 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA 1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1085 AATTACCATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGA 1144

Qy    1205 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACAAACG 1264
           ||||||||||||||||||||||||||||||||||||||||| | |||||||||||| | |
Db    1145 ATCAAGTTAAAACAGATGATATTGAAGCTTCAAGCACTGATCAGGGTACAACAACACAGG 1204

Qy    1265 GAACAACGACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGAAATACTA 1324
           |||||||||||||| |||||||||||||||||||||||||||||||||||||| | ||||
Db    1205 GAACAACGACAACAACTGATACATCTAGTGGTTCAACAGGTGCTGGAACAGGACAGACTA 1264

Qy    1325 CTAACACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1384
           || | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1265 CTCAGACTTCTCAAACAGTTTCTAATCCTACTTTAAATACTTATCGTAGTTTTGGAATTG 1324

Qy    1385 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAACGTTA 1444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1325 ATAGTAAACCAACTTCTGCAAACAAAATAGATGAAACTAATTGGGCAGATCCTAACGTTA 1384

Qy    1445 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1385 TTGAAGCAAGAATATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTA 1444

Qy    1505 ATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGTTCAA 1564
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1445 ATGCAGGAAACTTTATCCGAAACACAATTGGTGGTGTTGGTTTTACTTCAACAGGTTCAA 1504

Qy    1565 GAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAACCTT 1624
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1505 GAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAACGTCCAACTGGAAACTTCCAACCTT 1564

Qy    1625 TCTTATACGTATTTGGTTATTTAGGATACCAACAAACTAGAACAGGAACTTTCTGGTACG 1684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1565 TCTTATACGTATTTGGTTATTTAGGATACCAACAAACTAGAACAGGAACTTTCTGGTACG 1624

Qy    1685 GAACATATAAGCTTTTAAACAACAGCCCTTACGACGTATTAGATTCTCCAAGAGTAGGTA 1744
           ||||||||||||||||| | |||||||||||||||||||||||| || ||||||||||||
Db    1625 GAACATATAAGCTTTTACAGAACAGCCCTTACGACGTATTAGATGCTGCAAGAGTAGGTA 1684

Qy    1745 CTGAAACCAATCAATTTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATATCTAA 1804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1685 CTGAAACCAATCAATTTAGAAGAACTTCATTAACATACCCTGTTATGGGTGGATATCTAA 1744

Qy    1805 CTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1864
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1745 CTGAAGAAGGTGCTAGAAGTTTCTCTAATACTCCATATATAAGAGCACAAGGTGACACAC 1804

Qy    1865 CAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1924
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1805 CAGAAAGCCGAAGCATCTTCCAATCTGGCTATTCTGATAATACTTATGAGTACATTCAAT 1864

Qy    1925 CAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1984
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1865 CAGTTTTAGGATTTGATGGAATTAGAAATAACTTAAATGTTGGGGTTAAAGCATCAAGCT 1924

Qy    1985 TCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACATACT 2044
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1925 TCTTAAACTCAAATAGACCAAATCCAAACGGTCTAGAAATGATTGCTGCAACAACATACT 1984

Qy    2045 TAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAAACCAACAACCATTCGGAA 2104
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1985 TAAGATCACAAATTGGATTAGCTAGAACATCTGGATTACCAAACCAACAACCATTCGGAA 2044

Qy    2105 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCAGTTCTCATCAATTAAGAATATTA 2164
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2045 CAACTCACCAAGTTATTTCAGTATCACCTGGTGATCAGTTCTCATCAATTAAGAATATTA 2104

Qy    2165 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATAATAAAT 2224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||||
Db    2105 GAACAATCTTCCCTGGTAACCAGTTATGGTACTTCTTATTCACAAATGAAAATCAGAAAT 2164

Qy    2225 CTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCATTCA 2284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2165 CTAGTGTTTATACATTAAGATTAGCTGACTCAAGTAACCCTGATGCGTCAAGCTCATTCA 2224

Qy    2285 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGACAATT 2344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2225 GTCCAACAAGTTTAATTGACGTTAATGAAATTGGTGTAATCTTACCTTTATTAGACAATT 2284

Qy    2345 CATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGGTTCTC 2404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2285 CATTCTATACAGTAAATGCTGCTGGTAATGTTGCATTGTTCTCATCAAACCCTGGTTCTC 2344

Qy    2405 CTGGATCATATACTGCTGTAAATACATTTAATCAGAACTTATCTGATATTGCTTTTGAAG 2464
           ||||||||||||||||||||||||||||||||||| |  | |||||||||||||||||||
Db    2345 CTGGATCATATACTGCTGTAAATACATTTAATCAGCAGCTGTCTGATATTGCTTTTGAAG 2404

Qy    2465 GTTCTGGTGCTAAGTATACATCTGATTTCTGGGGAACAATCCAATTCAAACCCGATGAGT 2524
```

```
Db   2405  GTTCTGGTGCTAAGTAGTACATCTGATTTCTCGGGAACAATCCAATTCAAACCCGATGAGT  2464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2525  ACTTAATTCAAACAGTGGGTTCACTGTAGTCAGTCAGTGCTAGAAACTTCGTTACAAACCAAAGCT  2584
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   2465  ACTTAATTCAAAATGGGTTCACTAGTCACTGTCAGTGGCTAGAAACTTCGTTACCACAGGCAAAGCT  2524

Qy   2585  TCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTACCGTGTAGTGG  2644
           |||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db   2525  TCTTAAACAGTTAGTTGACTTCACTCCTGCTAATGCTGGTACTACCGTGTAGTGG  2584

Qy   2645  TTGATCCTGATGGTAATTAACAAACCAAAACCTACCTCTAAAAGTTCAGATCCAATACT  2704
           ||||||||||||||||||||||  ||| |||||||||||||||||||||||||||||||
Db   2585  TTGATCCTGATGGTCAGCTGACTAAACCCAAAAACCTACCTCTAAAAGTTCAGATCCAATACT  2644

Qy   2705  TAGATGGTAAGTATTATGATGCTAAATTAAAGAACAATAATTAGTAACATTCTTATA  2764
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   2645  TAGATGGTAAGTATTATGATGCTAAATTAAGAACAATAATTAGTAACATTCTTTATA  2704

Qy   2765  ACAACTTGCGCCCTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA  2824
           |||||||| |||||| ||||||||||||||| |||||||||||||||||||||||||||
Db   2705  ACAACTTTGGCGCGCCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTAGGTA  2764

Qy   2825  TTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAA  2884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2765  TTCTTGCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAAAGA  2824

Qy   2885  AATTACAAGACAAAGGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTG  2944
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2825  AATTACAAGACAAAGGGGTTCAAAACAACATTCAAAAAAGTTGATACCTTGACTGCTGCTG  2884

Qy   2945  TTGGTTCAGTTACAAGAAGATTATTACCCAAAACTGCTAACGTTAAGAAAAACCTGCTG  3004
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2885  TTGGTTCAGTTCAGTTTACAAGAAGATTATTACCCAAAACTGCTAACGTTAAGAAAAACCTGCTG  2944

Qy   3005  CTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACTGGTGCTGCTGCTCGTGCTGCTGTCTC  3064
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2945  CTTTAGGTGCTGGTAAATCTGGTGATAAGAAAACTGGTGCTGCTGCTCGTGCTGCTGTCTC  3004

Qy   3065  CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAGCTCACCAACTGGGCCTAAATCTG  3124
           ||||||||||||||||||||||||  |||||||||||||||||||||||| |||||||||
Db   3005  CAGCTAAACCATCTGCACCAAAAGTCAGCTCACCAGCTAGCTCACCAACTGCGCGCCTAAATCTG  3064

Qy   3125  GTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG  3184
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3065  GTGCGCCTACAAAAACCAACTGCTCCTAAGCCAGCTGCTCCAAAACCAACCGCTCCCAAAG  3124

Qy   3185  AATAA  3189
           |||||
Db   3125  AATAA  3129
```

RESULT 5
US-10-282-122A-27316
; Sequence 27316, Application US/10282122A
; Publication No. US2004002929129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Xu, H.
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 27316
; LENGTH: 3159
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-10-282-122A-27316

Query Match   3.9%;   Score 124.6; DB 13; Length 3159;
Best Local Similarity 63.5%;   Pred. No. 1.6e-15;
Matches 190; Conservative   0; Mismatches 109; Indels   0; Gaps   0;

```
Qy   2770  TTTGGCGCCTTACCTTCATGGGTAGTGCCTACAGCAATTGGTAGTACATTGGTATTCTT  2829
           |||||| |||||||||||||||||  |||| ||||||||||||||||||||||
Db   2797  TTTGCAGCACTCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG  2856

Qy   2830  GCAATTATGATCATCTTAGGATTAGCTAATCGGTATCGGTATTCCTTTAAGAGCTCAAAGAAAATTA  2889
           ||||| || ||| |||| ||| || || |||  |||||||||||||||||||||
Db   2857  TTTATCTTGTTAGTCTTAGGACTTGGGATTCCCAATGTACAGGGTAAGAAAACTC  2916

Qy   2890  CAAGACAAAGGGGTCAAAACAACACATTCAAAAAAGTTGATACCTTGACTGCTGCTGGT  2949
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2917  CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATACACTCACAACTGCTGCTGCGGT  2976

Qy   2950  TCAGTTTACAAGAAGATTATTACCCAAACTGCTAACGTTGGTGTGGTGAAAAAAAGCACCTGCTTTA  3009
           |||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db   2977  AGTGTGTACAAAAAGATTATTACCCAAACTGGTGATAAGAAACTGCGTCCTGCTCGTAAACCTGC  3036

Qy   3010  GGTGCTGGTAAATCTGGTGATAAGAAAACTGGTGCTGCTGCTCGTGCTCTCCAGC  3068
           |||||||||||||||||||||||||||  |||||||||||||
Db   3037  AAAGCTGCTAATCCTAGTGTTAAAAAAACTGCTGCTTTTTAAAAAAACCACCTGTTCAACC  3095
```

RESULT 6
US-10-205-220-1
; Sequence 1, Application US/10205220
; Publication No. US2003017066A1
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragmer
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193P1D1
; CURRENT APPLICATION NUMBER: US/10/205,220
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 08/545,528
; PRIOR FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 1
;   LENGTH: 580073
;   TYPE: DNA
;   ORGANISM: Mycoplasma genitalium
US-10-205-220-1

Query Match      3.9%;  Score 124.6; DB 15;  Length 580073;
Best Local Similarity   63.5%;  Pred. No. 2.9e-14;
Matches  190;  Conservative  0;  Mismatches 109;  Indels  0;  Gaps  0;

Qy     2770  TTTGGGCCCTTACCTTACCTTCCTCATGGGTAGTGCCTACCAGCAATTGGTAGTACCATTAGGTAGTCTT  2829
                  | ||   | ||| ||   ||     ||||| | |   ||||||||
Db   228703  TTTGCAGGACTCCCTGCATGGGTGATCCCTGTATCAGTAGGTTCTTCAGTTGGGATCTTG  228762

Qy     2830  GCAATTATGATCATCTTAGGATTAGCTATCGGTATTCCTTTAAGAGCTCAAAGAAAATTA  2889
                  | ||| | |  ||||| |
Db   228763  TTTATCTTGTTTAGTCTTTAGGACTGGGATTGGGATCCCAATGTACAGGGTAAGAAAACTC  228822

Qy     2890  CAAGACAAAGGGTTCAAAACAACAACATTCAAAAAAGTGATACCTTGACGTGCTGCTGGT  2949
                  |||||||| |
Db   228823  CAAGATGCATCGTTTGTTAATGTCTTTAAAAAGGTTGATCACTTCACAACTGCTGTCGGT  228882

Qy     2950  TCAGTTTACAAGAGAGATTATTACCCAAACTGCTAACGTTAGAAAAAAACCTGCTGCTTTA  3009
                  |||||||||||||| |
Db   228883  AGTGTGTACAAAAAGAGATTATTACCCAAACTGGTGTGTGGTGAAAAAAGCACCTAGTGCATTG  228942

Qy     3010  GGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGTAAACCTGCTAAAACCTGTTCCAGC  3068
                  ||||||
Db   228943  AAAGCTGCTAATCCTAGTGTTAAAAAAACCTGCTGCTTTTTAAAACCACCTGTTCAACC  229001
```

RESULT 7
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US2003008263 7A1
; GENERAL INFORMATION:
;   APPLICANT: ZWIEBEL, LAURENCE J.
;   TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
;   FILE REFERENCE: N82B9
;   CURRENT APPLICATION NUMBER: US/10/094,240
;   CURRENT FILING DATE: 2001-03-08
;   PRIOR APPLICATION NUMBER: 10/056,405
;   PRIOR FILING DATE: 2002-01-24
;   PRIOR APPLICATION NUMBER: 60/264,649
;   PRIOR FILING DATE: 2001-01-26
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match   2.0%; Score 62.6; DB 15; Length 4985;
Best Local Similarity   45.5%; Pred. No. 0.02;
Matches 333; Conservative 0; Mismatches 396; Indels 3; Gaps 3;

```
Qy    491  TTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATGATGGGACTAAATTTA  550
                 ||||           ||    |   |||
Db   2830  ATAATAATAATAATATATAAGACATAATAATAATATAATAATAATAATAATA  2771

Qy    551  ATTTTACAAACCAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCT  610
                 ||||
Db   2770  ATAATAATAATAATAATCATAATAAAATATTATTATTGATAATAATGTATAATAATGAAT  2711

Qy    611  TAC-CTAAAGATTTACCCCAGATTGGTATAACTTACATTCAAAGAAAGATCTTACCA  669
                 |||
Db   2710  AACAATAATAATTAATATAATAATAATAATAATTAATTAATAATAATAATAATAAT  2651

Qy    670  AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGTAGTGAACAAATGCT  729
                 | |
Db   2650  TATTATTGTTTATAATAATAATAATAATAATAATAATAATAATCATAATTATAAT  2591

Qy    730  GATGATGGGATGTTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAA  789
                 | |
Db   2590  AATAATAATATAATAATAAGTTGTCAAATAACCTCATTTAATATAATAATAATAAA  2531

Qy    790  ACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCAATGCCTGGT  849
                 ||||   | |    |||| |  |
Db   2530  AGTAATAATTATAATATAATAGTAATAATAGTATATATAATTAATAATATATATAAT  2471

Qy    850  GCAAACAATAGATACCATTCTCAATTGAATGTCAAGCATGACATAGAATTAAAACATCTTCCAA  909
                 | |
Db   2470  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT  2411

Qy    910  TTAGATGAAAAATTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAATATTACA  969
                 |||
Db   2410  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT  2351

Qy    970  AGATTAGCTACT  981
Db   2350  AAATTCTGTATT  2339
```

RESULT 8
US-10-056-405-10/c
; Sequence 10, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
;   APPLICANT: ZWIEBEL, LAURENCE J.
;   TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
;   TITLE OF INVENTION: USE THEREOF
;   FILE REFERENCE: N7841
;   CURRENT APPLICATION NUMBER: US/10/056,405
;   CURRENT FILING DATE: 2002-01-24
;   PRIOR APPLICATION NUMBER: 60/264,649
;   PRIOR FILING DATE: 2001-01-26
;   NUMBER OF SEQ ID NOS: 23
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 4985
;   TYPE: DNA
;   ORGANISM: Anopheles gambiae
US-10-056-405-10

Query Match   2.0%; Score 62.6; DB 15; Length 4985;
Best Local Similarity   45.5%; Pred. No. 0.02;
Matches 333; Conservative 0; Mismatches 396; Indels 3; Gaps 3;

```
Qy    253  ATCAAATTAGATAGTTTACTAAACCATTATATGGTTTAGATCTAAGTGATGATGTTGGT  312
                 || |   || |    ||| || |
Db   3070  ATAAAAAAAAAACAATAATAAACAATAATAAGAATAATAACATTATAAGAATAATAAGAA  3011

Qy    313  GGATACAAAGTAAAACAATAGTTTCA-GATTACACAACTAGCAGAAATAGATTTGATCA  371
                 | |
Db   3010  CAACAACAACAACAATAATAAACAAATAGTTTCA-GATTACACAACTAGCAATAATAAGAATAATAA  2951

Qy    372  AAGACAAACAAGAGCATATTATGCTCGTTGTTGGTAATGATGAAGCTTCAT-TTAA  430
                 | | |
Db   2950  TAACAATAGCAATAATAAGAATAGTTAATAATAATAATAATAATAATAATAATAA  2891

Qy    431  AAAGAATTAATACTAACTCAAATAGAATTGGTAATAGTAATAATATATAATATTAATATTA  490
                 ||| |     |    |||| |    ||||
Db   2890  AAAAAAATAATTATAATAATAAGAATAGTAATAGTAATAAATATAATATATTTAAAAAATAATATTA  2831
```

```
Qy   431  AAAGATTAATACTACTCAAATAGAATGGTAATAGAACAACAATTCTCAAGTTGTAA  490
          |||  ||||||||  |||||||                  ||||||||||||
Db  2890  AAAAAAAATAATTATTTAATAATAGAATAGTAATAGTAATAATAATAATAATATTA  2831

Qy   491  TTGGTGGTGTTGATAATCCAGCTCACGTAATTACTGATTACTGATGATGGGACTAAATTA  550
          |||||||||||||                      |||   ||||||||
Db  2830  ATAATAATAATATAATATAATAATAATAAGATAATATAATATAATATAATAATAATAATA  2771

Qy   551  ATTTTACAAACCAAACTCAAGGTGAAATTGTAATGACTTCATTTTAGATGCGCCAATCT  610
          |||                |||||||                   ||||||||
Db  2770  ATAATAATAATATAATCATAATAAATAAAATATTATTATTTATGATAATGTATAATAATGAAT  2711

Qy   611  TAC-CTAAAGATTTACACCCAGATTGGTATAACATTCAAAGAAGATCTTACCA  669
          |||     ||||||||                   |||||||||||
Db  2710  AACAATAATATAATAAATTCAAAAAATCAAAAAAAACCTAAAAAATAATAATAAT  2651

Qy   670  AATGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGAGTTAGTGGAACAAATGCT  729
          |||||||                          ||||||    |||||||
Db  2650  TATTTATTGTTATAATAATAATAATAATAATATAATAATAATAATCATAATTATAT  2591

Qy   730  GATGATGGGATGTTGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAA  789
          |||                      ||||||||                |||
Db  2590  AATAATAATAATATAATAATAATAATAATAATAATAATAATTTATATTATATTTATATTAA  2531

Qy   790  ACTAAAACCACTACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAATGCCTGGT  849
          |||                   ||||||||                |||||
Db  2530  AGTAATAATTATAATAATAATAGTAATAATAATAATAATAATTATAATAATAAT  2471

Qy   850  GCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATATGAATTAAAAACATCTTTCCAA  909
          |||               ||||||||                    |||||
Db  2470  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATTAATAATAAT  2411

Qy   910  TTAGATGAAAAATTTGTTTATCCAGAATGGACTGCGTTCTCGAAGAGAATAAAAAATATTACA  969
          |||||               |||||||                 |||||||
Db  2410  AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT  2351

Qy   970  AGATTAGCTACT  981
          |||||||||
Db  2350  AAATTCTGTATT  2339
```

RESULT 9
US-10-311-455-1798/c
; Sequence 1798, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;   TITLE OF INVENTION: cytosine methylation
;   FILE REFERENCE: 5013.1014
;   CURRENT APPLICATION NUMBER: US/10/311,455
;   CURRENT FILING DATE: 2002-12-16
;   PRIOR APPLICATION NUMBER: PCT/EP01/07537
;   PRIOR FILING DATE: 2001-07-02
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 2424
;   SEQ ID NO 1798
;   LENGTH: 8771
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1798

Query Match   1.9%; Score 60.8; DB 15; Length 8771;
Best Local Similarity 46.3%; Pred. No. 0.065;
Matches 200; Conservative   0; Mismatches 232; Indels   0; Gaps   0;

```
Qy   256  AAATTAGATAGTTTACTAAACCATTATATATGTTTAGATCTTAAGTGATGATGTGTGGA  315
          |||               |||||||              ||||||||
Db  5053  ATAATCTATACATACAATACAATAATATATTATTCAACTTTAAAAAAAAAAAACTTAATCTA  4994

Qy   316  TACAAAGTAAAACAAATAGTTCAGATTACCAACTAGCAGAAATAGATTTGATCAAAGA  375
          |||               |||||||             ||||||||
Db  4993  TACTACACAAAAAAAACCTTAAAAACATTATCAAAAATAAATAAACAATAACAAAA  4934

Qy   376  CAAACAAGAGCATATTATGCTCCTGTGGTTAATGAGAAGCTAACGTTCATTTAAAAAAGA  435
          |||               |||||||              ||||||||
Db  4933  AATACATACTATATATAATTCCACTTCTAAAAAAACTACCTAAAATAATCAAATTATAAAAAC  4874

Qy   436  ATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTGTGAATTGGT  495
          |||               |||||||              ||||||||
Db  4873  AAAAATAATATACATATAAATTCAAAAACCTAAAAAAAAAAATAATAATTATTTATTAAT  4814

Qy   496  GGTGTTGATAATCCAGCTCACGTAATTACTGATTACTGATGATGGGACTAAATTTAATTTT  555
          |||               |||||||              ||||||||
Db  4813  AAATAATAATAATATAATATTATACAAAAATATAAAACATAATTAATACCACTAAATTATACACA  4754

Qy   556  ACAAACCAAACTCAAGGTGAAATTGTAATGACTTCATTTTAGATGCGCCAATCTTACCT  615
          |||               |||||||              ||||||||
Db  4753  TACAAATAATCAAAAATAATTTTATATTATATTCATATTATATATTAATTTTTACCA  4694

Qy   616  AAAGATTTACACCCAGATTGGTATAACTTATCATTCAAAGAAAGATCTTACCCAAATGAC  675
          |||               |||||||              ||||||||
Db  4693  TAATACAAAAAAATAATAATAATAATAAAAAAAATAAAAAACTTAACGATTACATAACTTA  4634

Qy   676  GTCAACACTGCA  687
          |||||||
Db  4633  ACCAATACCTCA  4622
```

RESULT 10
US-09-894-273-1/c
; Sequence 1, Application US/09894273
; Publication No. US20040037847A1
; GENERAL INFORMATION:
;   APPLICANT: Kieff, Elliott D.
;   APPLICANT: Ballestas, Mary E.
;   APPLICANT: Kaye, Kenneth M.
;   TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
;   TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
;   FILE REFERENCE: 16412-10001R
;   CURRENT APPLICATION NUMBER: US/09/894,273
;   CURRENT FILING DATE: 2001-06-28
;   PRIOR APPLICATION NUMBER: US 60/109,422
;   PRIOR FILING DATE: 1998-11-19
;   NUMBER OF SEQ ID NOS: 3
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1
;   LENGTH: 3489
;   TYPE: DNA
;   ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-1

Query Match   1.9%; Score 60.2; DB 13; Length 3489;
Best Local Similarity 61.6%; Pred. No. 0.052;
Matches 114; Conservative   0; Mismatches 68; Indels   3; Gaps   1;

```
Qy  2998  CCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTCTAAACCT  3057
          |||||||           |||||||        ||||||||
Db  2236  CATCCTGCTGCTGCTCCATCCTGCTGCTGCTCATCCTGCTGCTCCTGCTGCTCATCCT  2177

Qy  3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTGCTGCTCCAGCTAACCACTGGGC--  3115
          ||||||           |||||||        ||||||||
Db  2176  GCTGCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCCTGCTGCTCCTGCTGCTCCT  2117

Qy  3116  -CTAAATCTGGTGCGGCCTACAAAACCAACTGCTCTGCTGCTCCTAAGCCAGCCAAAACCAACC  3174
          |||||||           |||||||        ||||||||
Db  2116  GCTCATCCTGCTGCTCCTGCTGCTCATCCTGCTGCTCATCCTGCTGCTCATCCTGCT  2057
```

Matches 209; Conservative   0; Mismatches 231; Indels   2; Gaps   1;

```
Qy   241  AATGGAACAATTATCAAATTATAGATAGTTTACTAAACCATTATATGGTTTAGATCTAAGT  300
               ||||     |||        |||   ||||| |||     |||   ||||||||||
Db 525047  ATTAAAAAAATAAACACAGAAAATTTAATACATACCATTCTAAAGTAATTTTATT 525106

Qy   301  GATGATTGTGTGATACAAGTAAACAATAAAACAATAGTTCAGATTACACAACTAGCAGAAAT  360
               |||||||||   |||     || |||||   |||   ||  |||||||||||||||||||
Db 525107  TTAAAAAAGAAGAAAAATATTATAAACTATTATCTCAAAAAGAGAATATAAAAAA 525166

Qy   361  AGATTTGATCAAGACAAGAGACATATATGCTCTGTGTTGGTTAATGATGAAGCTAAC  420
               |||   | |||| ||||    |||        ||||
Db 525167  GCAATTGTTAAAAATAGATCCTTCAATCTATAAAAAAATATAACAAATCAAGAAGCACACAA 525226

Qy   421  GTTCATTTAAAAGAGATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCT  480
                  |||        ||||         |||    ||||  ||      || |||
Db 525227  ATTTATTTCAAAAAAATCAAGATAATTTTTATATTCCAGAAAAAATTTAAAATTAAATTTT 525286

Qy   481  AAGTTTGTAATTGGTGTGTTGTTATATCCTCACGTAATAGATTTACTGATGATGGG  540
               ||| |||||||||||||||||  || ||||||||||  | |||   ||||||||
Db 525287  GTTGAATTAAAAACTGATAACTTAAAATACATTGCGAAAATAAAGAGATTTATGATTGG 525346

Qy   541  ACTAAATTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGAT  600
               ||    | |||  |  |||||||| |||     ||| ||||  | |||   ||||
Db 525347  TATATAAGAAAATATTAC--ACAATACTCAACTAAAGAAAAAGAGATATAGTATTATTC 525404

Qy   601  GCGGCCAATCTACCTAAAGATTTACACCCAGATTGGTATAACTATTACATTCAAAGAAAG  660
               ||||||||       |||     ||| |||| | |||    |||||||| |||||||
Db 525405  AAGTCAAAATAAACAACAGCCATATCGATATTATCTAGATTGCATAATACACCAGAAG 525464

Qy   661  ATCTTACCAAATGACGTCAACA  682
               ||   |    |||   |||||
Db 525465  ATTTTTCAAAAAATAGCTCAAGA 525486
```

RESULT 11
US-10-294-804-1/c
; Sequence 1, Application US/10294804
; Publication No. US20030113948A1
; GENERAL INFORMATION:
;   APPLICANT: Robertson, Erle S.
;   APPLICANT: Cotter, Murray A.
;   TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
;   TITLE OF INVENTION: to Genomic Host DNA
;   FILE REFERENCE: UM-03778
;   CURRENT APPLICATION NUMBER: US/10/294,804
;   CURRENT FILING DATE: 2002-11-14
;   PRIOR APPLICATION NUMBER: US/09/410,399
;   PRIOR FILING DATE: 1999-10-01
;   NUMBER OF SEQ ID NOS: 6
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 3489
;   TYPE: DNA
;   ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-10-294-804-1

Query Match   1.9%; Score 60.2; DB 15; Length 3489;
Best Local Similarity   61.6%; Pred. No. 0.052;
Matches 114; Conservative   0; Mismatches   68; Indels   3; Gaps   1;

```
Qy  2998  CCTGCTGCTTTAGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTGCTAAACCT  3057
               ||||||||||||| ||||||||||||||| ||||||||||||||||||||||||
Db  2236  CATCCTGCTGCTGCTCATCCTGCTCATCCTGCTGCTCACCAGCTAAACCAACTGGGC-- 2177

Qy  3058  GCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAACTGGGC--  3115
               |||||  ||||||||||||  ||   |||||  || ||||| |||   |||||||
Db  2176  GCTGCTGCTCATCCTGCTGCTCATCCTGCTCCTGCTCCTGCTGCTCCCAAAAACCAACC 2117

Qy  3116  -CTAAAATCTGGTGCGCCTACAAAAACAACTGCTCCTAAGCCAGCTGCTCCAAAAACCAACC  3174
                ||| |||||||  ||  ||||  || || || ||||  ||  ||  |||||||||||||
Db  2116  GCTCATCCTGCTCGTCCTGCTCCTGCTGCTCATCCTGCTGCTCTCATCCTGCT 2057

Qy  3175  GCTCC  3179
               |||
Db  2056  GCTGC 2052
```

RESULT 12
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
;   APPLICANT: SHIGENOBU, SHUJI
;   APPLICANT: WATANABE, HIDEMI
;   APPLICANT: HATTORI, MASAHIRA
;   APPLICANT: SAKAKI, YOSHIYUKI
;   TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;   FILE REFERENCE: 081356/0159
;   CURRENT APPLICATION NUMBER: US/09/790,988
;   CURRENT FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: JP2000-107160
;   PRIOR FILING DATE: 2000-04-07
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match   1.8%; Score 58.4; DB 9; Length 640681;
Best Local Similarity   47.3%; Pred. No. 2.3;

RESULT 13
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
;   APPLICANT: Epigenomics AG
;   TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
;   FILE REFERENCE: E01/1208/WO
;   CURRENT APPLICATION NUMBER: US/10/312,841
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

Query Match   1.7%; Score 53.6; DB 15; Length 3673778;
Best Local Similarity   47.1%; Pred. No. 59;
Matches 196; Conservative   0; Mismatches 219; Indels 1; Gaps   1;

```
Qy   801  TACTGATAATCAAAATCCTTCAACTTTTAATTCAGGAGCAACATGCCTGGTGCAAACAATAG  860
               ||||  ||||||||||||||||||| |||||||||||||||||  ||||||||||||
Db 98780  TAATTATCCAATTATTCCATCACCATTTAATAAAAAAAACTCTCCTTCCTTCATTAAAATTA 98721

Qy   861  AATACGAATTCTCAATTGAATGTCAAGCATAGAATTAAAAACATCTTCCAATTAGATGAAAA  920
               |||  | ||  |||||||  ||||||||||     |||||||  |||||||||||||
Db 98720  CTTTTATACTTTTTATAAAAAATCAATTAAAACATATTTTAATCATTTTAATTAAATTAACTC 98661

Qy   921  ATTTGTTTATCCAGAATGGACTGGTTCTGAAGAATGAAGATTAAAAAATATTACAAGATTAGCTAC  980
               ||   |||||||||||  |||||||||||   ||||||||| |||||||||||||||||
Db 98660  ATCTTCTTACATTAGCTTCTTAAAAAAACTTTCACCATATTCTAAATACATATTCTCATCAACA 98601
```

Qy   981 TGGAAGTTTGCCAAGCAACGAAAGATATTGGATTCTTGACATACCGGGACTCCACAAGT 1040
                  |||| || |||
Db   98600 AATATAATTTATAATAATAAATTATAAAATATTTTTCTCAATCTACGACTTATCTTCTCATT 98541

Qy   1041 TACTTTAAAAGAAGATTCAGTTAACGTAATTTCAGAGACTATACTTAAACTCAGTTAATTC 1100
                  ||| || |||||
Db   98540 TTCTCAATATTATCTTTAAAATACTTAAATTTAAATTTAACTAAACCCAATTTA-TC 98482

Qy   1101 TTTATCATTCATTGGTGATAGTATTTATATTTTTGGTACCTCTGAATTACCATCATTATG 1160
                  ||||||||| |||||||| ||
Db   98481 TCTCACTTTTTTTTCATCTATAAAATCATATTTTTTTATATTTTTATATAAAAATCTTTTACC 98422

Qy   1161 GTACTATTCATTCCCAACTAGATTATCTGATCTAACCGCTTTGAATCAAGTTAAAA 1216
                  ||||||| || |
Db   98421 TAACTACTCAATATCACAAAAATTTTCTCCTATATTTTCTTTAAAAATTTTATAA 98366

RESULT 14
US-10-373-667-3
; Sequence 3, Application US/10373667
; Publication No. US20030235841A1
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P. W.
; APPLICANT: Koob, Michael
; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/10/373,667
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/181,585
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA
; OTHER INFORMATION: comprising exons E, C, and A
US-10-373-667-3

Query Match   1.7%;   Score 53.2; DB 16; Length 1037;
Best Local Similarity 50.8%;   Pred. No. 0.77;
Matches 127; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   2930 CCTTGACTGCTGCTGGTTCAGTTACAAGAAGATTATACCCAAACTGCTAACGTTA 2989
                  ||||| | || |||||||||||||||||||
Db   605 CCTGGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTA 664

Qy   2990 AGAAAAAACCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTG 3049
                  ||| ||| |||
Db   665 CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724

Qy   3050 CTAAACTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
                  ||| |||||||||||||
Db   725 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784

Qy   3110 CTGGGGCCTAAAATCTGGTGCGCCTACAAAACCAACTGCTGCTGCTGCTCCAAAAC 3169
                  ||| |||||||||||||||
Db   785 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844

Qy   3170 CAACCGCTCC 3179
                  ||||||
Db   845 CTGCTGCTGC 854

RESULT 15
US-10-373-667-1
; Sequence 1, Application US/10373667
; Publication No. US20030235841A1
; GENERAL INFORMATION:
; APPLICANT: Ranum, Laura P. W.
; APPLICANT: Koob, Michael

; TITLE OF INVENTION: SPINOCEREBELLAR ATAXIA TYPE 8 AND METHODS OF DETECTION
; FILE REFERENCE: 11000900101
; CURRENT APPLICATION NUMBER: US/10/373,667
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US/09/181,585
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-373-667-1

Query Match   1.7%;   Score 53.2; DB 16; Length 1159;
Best Local Similarity 50.8%;   Pred. No. 0.82;
Matches 127; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   2930 CCTTGACTGCTGCTGGTTCAGTTACAAGAAGATTATACCCAAACTGCTAACGTTA 2989
                  ||||| | || |||||||||||||||||||
Db   417 CCTGGGGTCCTTCATGTTAGAAAACCTGGCTTTACTACTACTACTACTACTACTA 476

Qy   2990 AGAAAAAACCTGCTGCTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCTGCTG 3049
                  ||| ||| |||
Db   477 CTACTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536

Qy   3050 CTAAACTGCTGCTCCAGCTAAACCATCTGCACCAAAAGCTAGCTCACCAGCTAAACCAA 3109
                  ||| |||||||||||||
Db   537 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596

Qy   3110 CTGGGGCCTAAAATCTGGTGCGCCTACAAAACCAACTGCTGCTGCTGCTCCAAAAC 3169
                  ||| |||||||||||||||
Db   597 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656

Qy   3170 CAACCGCTCC 3179
                  ||||||
Db   657 CTGCTGCTGC 666

Search completed: May 6, 2004, 00:19:49
Job time : 1270.74 secs

```
                GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May  5, 2004, 10:50:12 ; Search time 7289.65 Seconds
                                   (without alignments)
                        13063.796 Million cell updates/sec

Title:          US-09-901-572A-2
Perfect score:  3189
Sequence:       1 atgaatatttctaaaaact........caaccgctcccaagaataa 3189

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 1491090276 residues

Total number of hits satisfying chosen parameters:   55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
            1: em_estba:*
            2: em_esthum:*
            3: em_estin:*
            4: em_estmu:*
            5: em_estov:*
            6: em_estpl:*
            7: em_estro:*
            8: em_htc:*
            9: gb_est1:*
           10: gb_est2:*
           11: gb_htc:*
           12: gb_est3:*
           13: gb_est4:*
           14: gb_est5:*
           15: em_estfun:*
           16: em_estom:*
           17: em_gss_hum:*
           18: em_gss_inv:*
           19: em_gss_pln:*
           20: em_gss_vrt:*
           21: em_gss_fun:*
           22: em_gss_mam:*
           23: em_gss_mus:*
           24: em_gss_pro:*
           25: em_gss_rod:*
           26: em_gss_phg:*
           27: em_gss_vrl:*
           28: gb_gss1:*
           29: gb_gss2:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	c	Score	%Match	Query Length	DB	ID	Description
1	c	73.6	2.3	1896	29	CG753083	CG753083 P048-1-C0
2	c	70.8	2.2	1101	29	CNS0039G	AL063921 Drosophil
3		67	2.1	1200	13	BX437758	BX437758 Drosophil
4		66.4	2.1	1101	29	CNS00EVL	AL069706 Drosophil
5		65.8	2.1	844	29	BX139987	BX139987 Danio rer
6	c	65.2	2.0	1101	29	CNS0039G	AL063921 Drosophil
7		64.6	2.0	829	29	BX173672	BX173672 Danio rer
8		64.2	2.0	773	29	CNS01VTG	AL169549 Tetraodon
9		63.8	2.0	1200	13	BX447454	BX447454
10	c	61.8	1.9	470	29	FR0018463	AL011359 F.rubripe
11		61.8	1.9	480	28	BZ643398	BZ643398 OGAOW57TC
12	c	61.8	1.9	843	28	BZ643413	BZ643413 OGAOW57TM
13	c	61.8	1.9	1225	29	CNS0161D	AL106171 Drosophil
14	c	61.6	1.9	735	29	CNS04NSM	AL299119 Tetraodon
15	c	60.6	1.9	641	28	AQ946120	AQ946120 Sheared D
16	c	60.2	1.9	1101	29	CNS00FVE	AL071298 Drosophil
17	c	60	1.9	941	13	BU126589	BU126589 60003315
18		60	1.9	1056	13	BX415058	BX415058
19		59.8	1.9	450	29	FR0025683	AL018519 F.rubripe
20		59.8	1.9	498	12	BJ396861	BJ396861
21		59.8	1.9	575	28	BZ780846	BZ780846 i120h11.g
22		59.8	1.9	615	28	BH766948	BH766948 BMBAC345F
23		59.8	1.9	994	13	BX414650	BX414650
24	c	59.2	1.9	619	29	FR0006944	Z90754 F.rubripes
25	c	59.2	1.9	1201	13	BX458623	BX458623
26		59		317	12	BJ328366	BJ328366
27		58.8	1.8	619	29	PR0013713	AL004959 P.rubripe
28		58.8	1.8	759	29	CNS06QXV	AL411257 T7 end of
29	c	58.8	1.8	766	14	CF348501	CF348501 AGENCOURT
30		58.8	1.8	813	29	BX143161	BX143161 Danio rer
31	c	58.8	1.8	836	28	BH491109	BH491109 BOGGY47TR
32		58.8	1.8	1007	29	CNS06X9S	AL419462 T3 end of
33		58.4	1.8	1126	13	BX446391	BX446391
34		58.2	1.8	332	13	C92788	C92788 C92788 Dict
35	c	58.2	1.8	641	12	BM181884	BM181884 fv51bl1.y
36	c	58.2	1.8	1350	29	CG744271	CG744271 P036-4-C0
37		58	1.8	257	29	FR0026883	AL019712 F.rubripe
38		58	1.8	939	29	CG900918	CG900918 ZMMBBb051
39	c	57.4	1.8	1101	29	CNS0100X	AL098379 Drosophil
40	c	57.4	1.8	1101	29	CNS00BSI	AL069797 Drosophil
41	c	57.2	1.8	427	28	BZ422321	BZ422321 id5zc10.b
42		57.2	1.8	649	12	BM160056	BM160056 EST562579
43	c	57.2	1.8	1101	29	CNS001FB	AL060732 Drosophil
44	c	56.8	1.8	639	29	CNS038CX	AL232458 Tetraodon
45	c	56.8	1.8	989	29	CNS02HA4	AL197365 Tetraodon

ALIGNMENTS

```
RESULT 1
CG753083/c
LOCUS       CG753083                1896 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P048-1-C01.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG753083
VERSION     CG753083.1  GI:37977199
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
ORGANISM    Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1896)
AUTHORS     Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE       An integrated physical and genetic map of the nematode Pristionchus
            pacificus
JOURNAL     Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE     22835951
PUBMED      12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
```

FEATURES　　Class: BAC ends.
　source　　Location/Qualifiers
　　　　　　1..1896
　　　　　　/organism="Pristionchus pacificus"
　　　　　　/mol_type="genomic DNA"
　　　　　　/strain="California"
　　　　　　/db_xref="taxon:54126"
　　　　　　/clone_lib="Ppa EcoRI BAC Library"
　　　　　　/note="The library was generated by a partial digest of
　　　　　　the genomic DNA with EcoRI and cloning into the BAC
　　　　　　vector."

ORIGIN

Query Match　　　2.3%;　Score 73.6;　DB 29;　Length 1896;
Best Local Similarity　42.8%;　Pred. No. 7.5e-05;
Matches 512;　Conservative　0;　Mismatches 675;　Indels　9;　Gaps　4;

Qy　315 ATACAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGATTTGATCAAAG 374
　　　　　|||　|||||　||||　|　|||||||||||　|||　|||　||||　||||||
Db　1786 AAAAAAATTAAAAAATAAAAAAATTAAAATATAAATAAAATAAAAAATTAAATAAATATATA 1727

Qy　375 ACAAACAAGAGCATATTATGCTCTGTTGGTTAGATGATGAAGCTAACGTTCATTTAAAA-A 433
　　　　|　|||||　||||||||||||　||||||　|||
Db　1726 AATATTATTAAATATATTAAAATATTNATAAAATATAAAAAATNTAAAATTATATANAAATA 1667

Qy　434 GAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTGTAATTG 493
　　　　|　||||||||||||||||||||||||
Db　1666 TAATTAAAAATTAAAATAAATATATTANNAAAAAAAAAATATTTTTNAATTTATTATATA 1607

Qy　494 GTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAAATTTAATT 553
　　　　|||||||||||||||||||||||
Db　1606 TAATAATAATTATTTAAAATATTATTAAAATAATAAATATATTTTTTATATTTTTAT 1547

Qy　554 TTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTGATGCGCCAATCTTAC 613
　　　　||||||||||||||
Db　1546 TTTTAATATATAATAATTTTTAATATTTATTAAAAATAATATATATANATAAAAATTAT 1487

Qy　614 CTAAAGATTTACACCCAGATGGTATAACTTATACATTCAAAGAAAGATCTT--ACCAAA 671
　　　　|||||||||||||||
Db　1486 ATTTTTTTTTATTTTTTTTATNTAAAATATTTATAAAATAAAATTATNAAATAAA 1427

Qy　672 TGACGTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGGTAGTCGAACAAATGCTGA 731
　　　　|||||||　|||||||　||||||||||||||||||
Db　1426 AAAAAATTATAAANTAAATTTTATTTTTACTTTAAAATATTTAAAATATATTNANNAA 1367

Qy　732 TGATGGGATGTTTGATTGTGGGAAATGGTCAAATAACTAATACACAGATCCTATTGCTCAAAC 791
　　　　|||||||　|||||||||　|
Db　1366 TATAAAAATATATAAAAAATATNATAAATAAAATTAAAATTATATTTTTTAAAATTAT 1307

Qy　792 TAAAACCACTACTGATAATCAAAATCCTCAACTTTAATTCAGGAGCAATGCCTGGTGC 851
　　　　|||||||||||||||
Db　1306 TAAATTAAAATTATAAAAATAAAAATAAAAAAAAAATTTATATAAA---TAATAATAATATNA 1251

Qy　852 AAACAATAGATACGATTCTCAAATGAAATGTCAAGCCATAGAATTAAAACATCTTTCCAATT 911
　　　　||　|||||||　|||||||||||||||
Db　1250 AAAAAAAATAAAAAAAAAATAATAATATAATTAATAATTATATATATAAAAATAAATAAT 1191

Qy　912 AGATGAAAAATTTGTTATCCAGAATGGACTGGTTCTGAAGAAGAAATAATATTACCAAG 971
　　　　|||||||||||
Db　1190 TTAATAAAATTAAATTAATTATTATAAATAATAAATAAAAAAATAAATATAAAAAAAAAA 1131

Qy　972 ATTAGCTACTGGAAGTTTGCCAACGAAAGATATATCTTGACATACCCGGGAC 1031
　　　　||||||||||
Db　1130 ATAAATATTATTAAAATTAAAAAATAAAAATNATTNATTTTATTNTTAAAATATATAAATAT 1071

Qy　1032 TCCAAGTTACTTTAAAAGGAAGAGATTCAGTTAACGTATTTTCAAGACTAACTTAAACTC 1091
　　　　||||||||||
Db　1070 TTATTTATTTATATATATTAATTAAAATAATATAATATTNTTTATATATTAATAAAAAATTTA 1011

Qy　1092 AGTTAATTCTTTATCATTCATTGGTGATGAATAGTAATTAATTATTTTTTGGTACCTTCTCGAATTACC 1151
　　　　|||||||||　|||||||
Db　1010 TAATAAATTTTTTAAATAAATTAAATAAATANATAATNTATTTTATTTATTTTATTTATTAAAAAATATAT 951

RESULT 2
CNS0039G/c
LOCUS　　CNS0039G　　　1101 bp　　DNA　　　linear　　GSS 03-JUN-1999
DEFINITION　Drosophila melanogaster genome survey sequence TET3 end of BAC #
　　　　　BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
　　　　　fly),genomic survey sequence.
ACCESSION　AL063921
VERSION　　AL063921.1　GI:4941778
KEYWORDS　GSS.
SOURCE　　Drosophila melanogaster (fruit fly)
　ORGANISM　Drosophila melanogaster
　　　　　Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
　　　　　Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
　　　　　Ephydroidea; Drosophilidae; Drosophila.
REFERENCE　1　(bases 1 to 1101)
　AUTHORS　Genoscope.
　TITLE　　Direct Submission
　JOURNAL　Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
　　　　　BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
　　　　　- Web : www.genoscope.cns.fr)
COMMENT　　Determination of this BAC-end sequence was carried out as part of a
　　　　　collaboration with the Berkeley Drosophila Genome Project (BDGP).
　　　　　The BDGP is constructing a physical map of the Drosophila
　　　　　melanogaster genome using these BACs. For further information
　　　　　please see http://www.fruitfly.org The BDGP Drosophila
　　　　　melanogaster BAC library was prepared by Kazutoyo Osoegawa and
　　　　　Aaron Mammoser in Pieter de Jong's laboratory in the Department of
　　　　　Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
　　　　　NY. The library is named RPCI-98 and was constructed by partial
　　　　　EcoRI digestion of Drosophila DNA provided by the BDGP from the
　　　　　isogenic strain y2; cn bw sp, the same strain used for the BDGP's
　　　　　P1 and EST libraries. A more detailed description of the library
　　　　　and how to order individual BAC clones, the entire library, or
　　　　　filters for hybridization from the BACPAC Resource Center can be
　　　　　found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES　　Location/Qualifiers
　source　　1..1101
　　　　　　/organism="Drosophila melanogaster"
　　　　　　/mol_type="genomic DNA"
　　　　　　/db_xref="taxon:7227"
　　　　　　/clone="BACR08K10"
　　　　　　/clone_lib="RPCI-98"
　　　　　　/note="end : TET3"

ORIGIN

Query Match　　　2.2%;　Score 70.8;　DB 29;　Length 1101;
Best Local Similarity　16.3%;　Pred. No. 0.00028;
Matches 227;　Conservative 255;　Mismatches 94;　Indels　0;　Gaps　0;

Qy　1152 ATCATTATGGTACTATTCATTCCCAACTAGATTATCTGATCTAACCGTTTGAATCAAGT 1211
　　　　||||||||||||
Db　950 TAAAAAAATTATTTAATTTAATTAAAAANAATTAAAATTATTAAAATTATTTATATATA 891

Qy　1212 TAAAACAGATGATATTGAAGCTTCAAGCACTGATAACGGTACAACAACAAACGGAACAAC 1271
　　　　||||||||||
Db　890 AAAAAAATAAATAAATTAAAAAAAAAAATANAAAATAAAATATATTAATTATA 831

Qy　1272 GACAACAGCTGATACATCTAGTGGTTCAACAGGTGCTGGAACACAGGAAATACTACT--AAC 1329
　　　　||||||||||
Db　830 TATAATTTATNATNATNAAAAAATTATAAAATTAAAATTAAATATAAATAATTTTTATTATTTAAAT 771

Qy　1330 ACTTCTCAAACAGTTCTAATCCTACTTTAAATACTTAAATCATCGTAGTGTTTGGAATTGATAGT 1389
　　　　|||||||||　|||||||
Db　770 ATTATTAAATATATAAATAATAATAAAATTAATAAATTTATAATAATAAAAAAATAAAAATAAA 711

Qy　1390 AAACCAACTTCTGCAAACAAATAGATGAAACTAATTGGGCAGATCCTAACGTTATTGAA 1449
　　　　|||||||||
Db　710 AATTAAAAATATATAAATATATAATATTTTTTTTATTTATTATTTTATAATAATAATAATAATATAA 651

Qy　1450 GCAAGAATATGCTGAATACAGATTAGGTATTCAAAATGAAATTCCAATAACTAA 1505
　　　　||||||||||　||||||
Db　650 TTATNNAAATATAATAATTATTATTTATAAAATTATNAATTTATTAAATATATATATATAAAATATA 595

```
Qy      7  ATTTCTAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGAGCTCTT  66
              |:||| |::| ::|  |::  ::: |||:: :||::  :|:|| |:::  |:::  :
Db   1051  AGTWTWTATWWWWWWWATWDTWKDKWWWWATAAKTDTAWTWWRTAWRADWAGRDRGAGKR  992

Qy     67  GGTTCTGCAAGCTTTGCTTTAAGCAATCAGATAAGAGTAACGATAACACGCAATTAGTT  126
              :::  ||  |  :: ::|:  |:|:::::|::::::::::::|:::::::
Db    991  DRDAATDADGAGRRDGRKRKDKDKDRKDGDDDKKGGKKKKAAKAKWATKWWDDWDWDKD  932

Qy    127  AATCAAGCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGT  186
              :  :||:  :|||: |::|| ||||||:| ::|:| |::|:::  ||||:::|||
Db    931  WKWDGJAKDRKADDDDGAGDKDDDGKGKDADDDTDGTKDDDDWDKAKGTWGDATW  872

Qy    187  TCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAATGGA  246
              |:|| |::||:|::| ::|:||||::|:|| |::|::||:|:|:||:||::|||
Db    871  AWAATDWWWWGWADADWWTWDLAADDWWADDRWDAWAWKWDDAWAWGARTADRRDWGDRA  812

Qy    247  ACAATTATCAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCAAGTCAAGTGATGAT  306
              ::  :||||:|: |:||:|:: |: |||:|||||:|||:||:|::||::
Db    811  GKRRGGARKRRDRKRADDKRADDRDDAATWTWTTWTTTTRDTDDWKWKTDTWTRWAADRT  752

Qy    307  TGTGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGATTT  366
              |:::::|:|:  ||||:::|:||:::::|:::| ||::|::::| |:|||:|:
Db    751  WDRDDDDDRDRAGTAGRKWRRTWKRMKRRDTRWDDADADDTARDDRRRRGDDGADAGKG  692

Qy    367  GATCAAAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTCAT  426
              :||:||:|:||:|:|||: |::| : ::|| ::: :|:::| |::|:::||| ||
Db    691  KKTGRKRRRRDRATWDRTDAWWADAAWWTTTDTDTDWDKDRDRRRKGARRRRRTTARAAW  632

Qy    427  TTAAAAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTT  486
              ||:|||:|||  ::: |: |:|: |:|::|::|::|::|::|::|||||:
Db    631  DWWTWKAWDWAKWDWKTRADRWDRWAADTWTDARKADRDWAKARAWRARRDRARAARADR  572

Qy    487  GTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTCAGTAATTTACTGATGATGGGACTAAA  546
              ||:|||:||| ||::|:||:|:|||:||| ||:|||::: ||:|:|:
Db    571  RWTTKGKTTTATWTTWAARAAWWAWAWATTTATWTTWTTTTTWTTTTTWTTAA  512

Qy    547  TTTAATTTACAACAACCAAACTCAAGGTGAAATTGTT  582
              :::||||: ||: |:  |:||||:|||||:
Db    511  WWWAAWWTATWAAWTAAAWAAAAAAAAWAATTTTTT  476
```

RESULT 3
BX437758

```
LOCUS       BX437758            1200 bp    mRNA         linear     EST 15-MAY-2003
DEFINITION  BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
            5-PRIME, mRNA sequence.
ACCESSION   BX437758
VERSION     BX437758.1  GI:30773605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1200)
AUTHORS     Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0CAP008CA01QP1.
FEATURES             Location/Qualifiers
     source          1..1200
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0CAP008YB01"
                     /tissue_type="THYMUS"
```

/clone_lib="Homo sapiens THYMUS"
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

```
Query Match                2.1%;  Score 67; DB 13;  Length 1200;
Best Local Similarity     30.9%;  Pred. No. 0.0018;
Matches   283;  Conservative 181;  Mismatches 440;  Indels 13;  Gaps  4;

Qy     13  AAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTGGAGCTCTTGGTTCT  72
              |||||| |  |||||:||| ::|   ::|  |::  |: :||::||:
Db    266  AAAAAAAAAAAAWAAATWWWAAAAATWDTGATKADDATTGGGGGDWTTWWWTWTWTAA  325

Qy     73  GCAAGCTTTGGCTTTTTTTTWWAAAAARATAKRGGRGGADWATAWWTWTTTATWWWWDDDDAK  132
              |: :|:  |   ::| ::: :::::|:::::::||::::::::
Db    326  RRRDKTTTTTTTTWWAAAAARATAKRGGRGGADWATAWWTWTTTATWWWWDDDDAK  385

Qy    133  GCAAGAACGCTAGATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGTTCGTTG  192
              :||:|::  |:|:|:::|:|::::::|:: |:::||:||:::||::::::|
Db    386  AWADAGAATWTTTTTWTTWRAGRARKRDGTTDTRAADWWWTWAAWAWWWDKRKTTTTT  445

Qy    193  TTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAATGGAACAATT  252
              ::::|: |:|::|||::::::|||||:::|||||:|:::|:::|:::::|:|:::
Db    446  TTKKRRAAAATTTT----TTTTTTTTWTTWAKAAAAARAATAWTTWWRAAAAAAAWTW  500

Qy    253  ATCAAATTAGATAGTTTACTAAACCATTATATGTTTAGATCTTCAAGTGATGATGTGGT  312
              |:|:: |: :::: ::|| |  |  :||:||:::|:::|||:::::|:|:||
Db    501  WWAARGRARGATTTTTTTAAWWWGGGAGRTAWWTWTTTTTTTTTTTAWWWATWT  560

Qy    313  GGATACAAAGTAAAACAAATAG--TTTCAGATTACACAACTAGCAGAAATAGATTTGATCA  371
              ||:| :|: |:|:::|:::|||   ::| ||::|:::::||:: |:| |:::::|:
Db    561  ATATTTTTTWTWANAAAAAASAGAWKAATWTTTTKKTTATWWKRAGGATWTTTTTTTWTW  620

Qy    372  AAGACAAACAAGAGCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTCATTTAAA  431
              ||: |:::|:||| ::| :::  ::::::|: |::::::||::||:::|::| ::
Db    621  AARGDWKWTWTWTTTTTTTTTTTTTWTGTTTWWWRDGARRTAKATAKWWTTWTTWWT  680

Qy    432  AAGAATTAATACTAACTCAAATAGAATTGGTAATAGAACAACAATTCTAAGTTTGTAAT  491
              |:||: :|:|::|::| :|::| |::||::|::|::|:::||:|::||::|:: |:|
Db    681  AWAWAGGAARKRRATWTWTWTWAAAWAWWAGAAAAARGAGAGRAAWTTTTTTATKGAGRRG  740

Qy    492  TGGTGGTGTTGATAATCCAGCTCACGTAATTCAGTAATTTACTGATGATGGGACTAAATTTAA  551
              ||:|| ||::|:|||:|:| ::: |::|:: ||:|::::||::||
Db    741  ATWTTAWAWTWTATRARGADTTTTTTTAFAWWTATTTTTWWAWAGATDKAAAAAWWW  800

Qy    552  TTTACAAACCAAACTCAAGGTGAAATTGTTAATGTTAATGTGAAAAAAAAAAAAAAWRARWAATATATWTTT  611
              ::|:|:|:|:|||::|||||: :::|::
Db    801  WTTWTAAAAAAWWAATTWWWAGARAAAKATWTWAAWAAAKATWTWAAWAAAAAAAWRARWAATATATWTTT  860

Qy    612  ACCTA-AAGATTTACACCCCAGATTGGTATAACTTATACATTCAAAGAAGATCTTACCCA  670
              |:|:| ||: ||||:: :||::|:::|:::|::|::|::::::||||| :| ||
Db    861  ATATATAAKARAARARAAATAAAATAARARRWKGAAAAARAAAATAAAAAWWTTWTATAAA  920

Qy    671  ATGACGTCAACACTGCAGTTGTTCCTGGCCAGTAGGTAGTGTTGAACAAATGCTG  730
              |:||::|:|:|:|:|::|::|:::|:||::|:||||| :||:|:||||:|::
Db    921  AAAWTTTWAWWRAAAAWWDTTTTATWTAWTAWTATAGAWRAATARDATWRAA  980

Qy    731  ATGATGGGATGTTGATTGTTGGGAATGGTCAAATAACTAACTACAGATCCTATTGCTCAA  790
              |:|||::|:|||:||:||::|:||||::|::: ||::|::::|:|||
Db    981  AAAAWTTTTTTWTTWARAAAAWWDTTTTATWTATWTAWTATAGAWRAATARDATWRAA  1040

Qy    791  CT-----AAAACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCAATGC  844
                    |:||::|::|||::|| :|::::|||:::| :|::|| :||::::|:
Db   1041  TTATWADARRAAGAWATAXDAGAGATAAAWTTTTTTTTTTTTWTTWTAAAAAAAAGAAAA  1100

Qy    845  CTGGTGCAAACAATAGATACGATTCTCAATTGAATGTCAAGCATAGAATTAAAACATCTT  904
              ||:::||:|:|:|:|:|:|:|:||::|:|:||:|||:|::|:::|:|::| :||
Db   1101  TWTTARAAAGAWWATAAAAARAAAWATAAAAAAAAAAAAGAAKADAWAAGAAATAAA  1160

Qy    905  TCCAATTAGATGAAAA  921
```

Db   1161 AARAAAAAAGWRAARAW 1177

RESULT 4
CNS00EVL
LOCUS   CNS00EVL   1101 bp   DNA   linear   GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069706
VERSION   AL069706.1 GI:4949949
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR29B23"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
ORIGIN

Query Match   2.1%; Score 66.4; DB 29; Length 1101;
Best Local Similarity   35.1%; Pred. No. 0.0024;
Matches 224; Conservative 104; Mismatches 300; Indels 10; Gaps 3;

Qy   4 AATATTTCTAAAAAAAACTTAAAAGTTATACATTGATAGGTGGATTAGCTGTATTTGGAGCT 63
             : :|: |||||||||| | | |:||| || :| ::|||||| :
Db   471 WAATWTWWAAAWAAAWTTATWAATWAATWAAAAWAWWWAWTTTWWWWWWTWATTWTTTWAWW 530

Qy   64 CTTGGGTTCTGCGAGCTTTGGCTTTAAGCAATCAGATAAGAGATAACGATAACACGCAATTA 123
             :| | | |||||||| || | :: || :|| :: |||| :
Db   531 TWTAWTAAAAAAAAAAAWATAATTTAAAWWAATAWATTAAWAAWTATATTAATW 590

Qy   124 GTTAATCAAGCAAGAACGCTAGATGCTAATTCTCGTTAGACTTGCCAGGTCTTGGACAAAAT 183
             ::: || | || ||| ::| || | :|
Db   591 TATAAATWTWATTAATAAAAAAATATTTTTTTWATAAAATTTTATTAATTAATTAWT 650

Qy   184 GGTTCGTTCGTTGTTCAATACAGTTCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAAT 243
             :| |: || | :| :: | | :|| :
Db   651 TATTAAATAAWTWATTWWTAATTAAATAATTTWAAATAWAAAAAAAAAAAAWATWA 710

Qy   244 GGAACAATTATCAAATTAGATAGTTTTACTAAACCATTATATGTTTAGATCTAAGTGAT 303
             :||| |:|||: || | :::|:| :: |
Db   711 AWAATWATAWATAWT TAAAAWAATTAAAAWAAAWAAWWATWAWWATAATATATWWTTW 770

Db   304 GATTGTGGTGGATACAAAGTAAAACAAATAGTTTCAGATTACACAACTAGCAGAAATAGA 363
             :|: :::||||: ||| : || | :|| :|| ::|| :|| ||:|:|| |||
Qy   771 TAWWNATWWAWWWTATWAWATAWAAATAWAAWAWAWATAAATAWATAWATW--AAAWAWAWA 828

Qy   364 TTTGATCAAAGACAAACAAGAGCATATTATGCTCTCGTTGGTTAATGATGAAGCTAACGTT 423
             ||| :|:| : |:| | :: ||| | :|| ::||| :| |:
Db   829 TAWWATWATATAWWAATAWWAAAAAAAATWTAATAWCWTWAATWATAAWAAAAAATAWAWTTWW 888

Qy   424 CATTTAAAAGAATAATAACTCAAATAGAATTGGTATAGAAACAACAACAATTCTAAG 483
             || :|||| | ||| || |||| :|| || || || : :
Db   889 TWTTTTWWAWWATATAAWAWATAWAAAWAAWAAA----AAAAAAAATAWAWWWTWATA 943

Qy   484 TTTGTAATTGGTGGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACT 543
             ||| | |:|::|:|:|: : || | :: |:|| | | ||| :| :|
Db   944 TTWTTATTAAAWTWATWWAATTWWAWTTWWAWWTWATAWTTWTATWTATATTWWT---AWT 1000

Qy   544 AAATTTAATTTACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTAGATGATGCG 603
             |||| ||:| | || :| :::: |||| ||||||| ||| | | |::| :|| | |
Db   1001 AWATATATTTATTATTAAWWTATATTTTWAAAWTATAATATWATWTATAWWTWAWATATAWWAA 1060

Qy   604 CCAATCTTACCTAAAGATTTACACCCAGATTGGTATAA 641
             |||| ||| |: |||:|:|||| ||:|| |||:|
Db   1061 TTAWTTATATATTWTAAWATAWWATAWWAAAAAAWAWATAW 1098

RESULT 5
BX139987
LOCUS   BX139987   844 bp   DNA   linear   GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-99E7, genomic survey sequence.
ACCESSION   BX139987
VERSION   BX139987.1 GI:27971314
KEYWORDS   GSS.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 844)
  AUTHORS   Humphray,S.J., Huckle,B. and Durham,J.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT   This sequence was generated from the T7 end of BAC 99E7. 99E7 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES   Location/Qualifiers
     source   1..844
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-99E7"
                     /tissue_type="Testis"
                     /note="vector pindigoBAC-536"
ORIGIN

Query Match   2.1%; Score 65.8; DB 29; Length 844;
Best Local Similarity   45.2%; Pred. No. 0.0031;
Matches 241; Conservative   0; Mismatches 292; Indels   0; Gaps   0;

Qy   47 TAGCTGTATTGGAGCTCTGGTTCTGCGAGCTTTGGCTTCAAGCAATCAGATAAGAGTA 106
             || :||| |:|| || | || :| ||| |:| :| | |:| || |:|
Db   312 TTGCAGTCATTATTGCTTTTATTACTTTGATCATTTATTAATAATAATATAATAATAATAATA 371

Qy   107 ACGATAACACGCAATTAGTTAATCAGCAAGAACGCTAGATAATAATTCTGTTAGACTTG 166
             ||||| | ||| |:|| | |: | |: :| |||| ||||| || :
Db   372 ATGATAATAATAACGATAATAATTTWAAATAWAATAATATTGATAATAATGATAATAATA 431

Qy   167 CAGGTCTTGGACAAAATGGTTCGTTGTTGTTCATCACAGATCTAAGTGAT 226
             || |:|||||: || |:| |: | ::|| :|: :|| |: :
Db   432 ATAATGATAATAATAATATTGATAATAATTGATAATAATGATAATAATAATA 491

```
Qy   227 TTATAACGAGCAGCTAATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATG 286
             ||||||||  ||||  ||||||        ||||     |||||||  ||||
Db   324 AAAACAAATAGTTTCAGATTCAACTAGCAGAAATAGATAGTTTGATCAAAGACAAACAAG 383

Qy   492 ATAATATATTGATAATATATTGATGATAATAATAATAATAATAATAATGATAATATTGATAA 551
             :: |||  |||                |:::                 ||
Db   419 AAWAAAATAATTWAAAWAWAAAAAAATTWAAAAWAAAAAAWTAWWTTTAWTWAAAAAAAA 478

Qy   287 GTTTAGATCTAACTAACTGGTGATGATTGGTGGTGGATCACAAAGTAAAAACAAATAGTTTCAGATTACA 346
             |||||  ||| ::||  ||||  |||| ||   |:   :   |||||||
Db   384 AGCATTATGCTCTGTTGGGTTAATGATGAAGCTAACGTCATTTAAAAAGAATTAATAC 443

Qy   552 ATAATGATAATATATTGATAATATATTGATGATAATAATAATAATATGATGATAATATGATGATA 611
                 || ||          ||||||  |::|| |||
Db   479 AAAATTWTTTTTTTTTWTTAWTTWATAWTWTWWWTTAAAWAAAWAAAWAAAAWA 538

Qy   347 CAACTAGCAGAAATAGAGATTGATCAAAGACAAACAAGAGCATATTATGCTCTGTTGTTA 406
             ||||| |||||||||                 :|| |||||||| | ||||
Db   544 TAACTCAAATAGAATTGGTATAGAAGACAACACAATCTAAGTTTGTAATTGGTGGTGTTGA 597

Qy   612 ATAATGATAATAATAACAATAATAATAATGATAATAATAATAATGATGATAATAATAATA 671
             ||| | |||                         |||||
Db   539 TAAATWTWWTW-WTTYTTWAAWATAAAMCMAAWYHTYTYTYTYTYTYTYMTWHYHTM 597

Qy   407 ATGATGAAGCTAACGTTCATTTAAAAAGAATAAATACTCAAATAGAATTGGTAATA 466
             |||  |  ::||||  || ||              : |||| ||
Db   504 TAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAAATTTAATTTTACAAACCA 563

Qy   672 ATAATAATAATAATGATGATAATAATAATAATGATCAATAATAATGACAATAATAATAATG 731
             || :  :    |||||||||||                       :
Db   598 YTHAWAHTTWYHWYHTYAMWHWMTWHWTMWAWWHWTYTAAYYYTCMYYHYMGWHHA 657

Qy   467 GAAACACAATTCTAAGTTTGTGTAATTGGTGGTGTTGGTAATCCAGCTCACGTAATTAGAT 526
             |||||| :|  ||||    ||:||| | || |||||:|  ||||
Db   564 AACTCAAGGTGAAATTGTTAATGACTTCATTTAGATGACTCTTACCTAAAGATT 623

Qy   732 ACAATAATAATAATAATAATGATAATAATAATGATACTAATAATAATAATG 791
                ||| | ||||||:| |||        ||||||||||||
Db   658 HAHAAAWTTHTWWTHAYHWATYHYYYMYCAMMCMCTHTCHHCYYYHHYTAHHTHTHH 717

Qy   527 TTACTGATGATGGGACTAAAATTTACAAAACCAAAACTCAAGGTGAAATT 579
             |||||||||| :|||            ||||||
Db   624 ACACCCAGATTGGTATAACTATACATTCAAAGAAAGATCTTACCAAATGACGTCAACAC 683

Qy   792 ATAATAATAATGATAATAATAATAATGATGATAATAATAATGATGATATAT 844
             :::| |  :  :  || ||::  | ||
Db   718 WYAHYYMWYYMWAYYWMYCTACTYHYHHHHHYHWAYHTTWYAWAHAMWMWHHAHYAAAAA 777

Qy   684 TGCAGTTGTTCCTTGCCCAGTAGGTAGAGTTAGTGGAACAAATGCTGATGAGTGGGATGTT 743
             : ||||| ||||     | |:|||| ||||  ||| : || |||| :  ||||
Db   778 WAAWATTHHYHTTHYMHHTYMHYTYMYTCCYMCTYHCWHYYHTAYTCWTWHHWWWTWT 837

Qy   744 TGATTGTGGGAATGGTCAAATAACTAATACAGATCCTATTGCTCAAACTAAAACCACTAC 803
             | |||    |:   || ||| ||||:: |||||
Db   744 HWYHHTWWHHTTHWAWWHTHTWCWWWHATTWTWATHCWACMTWHHWMEMHHHHMAC 897

Qy   804 TGATAATCAAAATCCTCAACTTTAATTCAGGAGGCAATGCCTGGTGCAAACAATAGATA 863
             |   ||  ||:: ||  |||||  ||| :| |    ::  | : :|
Db   898 HAHHHTHMCMCHHHMCTCHHHHTMYHMTCHWMMWMHWHHWWMATWMTMTMWMCC 957

Qy   864 CGATTCTCAATTGAATGTCAAGCATAGAATTAAAAACATCTTTCCAATTAGATGAAAAATT 923
             :::|| :    |  :  ||||  ||||||  |||
Db   958 MMHHHCHMYHMHMHMYMYCCHYYCTCHTHATTHYHYMCTCYHYCTWHTYWTAYWWAWTAHA 1017

Qy   924 TGTTTATCCAGAATGGACTGGTTCTGAAGAAGAATAAAAATATTACAAGATTAGCTACTGG 983
             : |||: ::||||  | :: ||  ::|:
Db   1018 MTTATWWWMWMEWWAEWATWWWWWWWATAWWACTCHHWTWYHTHCTWYYHHTYHMWAWW 1077

Qy   984 AAGTTTGCCAAGCAAC 999
             :||  ::||: ::|:
Db   1078 MAWHWHHMYAHYHWAH 1093
```

RESULT 7
BX173672/c
LOCUS   BX173672   829 bp   DNA   linear   GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-150M6, genomic survey sequence.
ACCESSION   BX173672
VERSION   BX173672.1   GI:28005377
KEYWORDS   GSS.
SOURCE   Danio rerio (zebrafish)
  ORGANISM Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 829)
  AUTHORS   Humphray,S.J., Huckle,B. and Durham,J.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT   This sequence was generated from the SP6 end of BAC 150M6. 150M6 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES   Location/Qualifiers

```
Qy   227 TTATAACGAGCAGCTAATGGAACAATTATCAAATTAGATAGTTTACTAAACCATTATATG 286
```

RESULT 6
CNS0039G
LOCUS   CNS0039G   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL063921
VERSION   AL063921.1   GI:4941778
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp; the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR08K10"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
ORIGIN
Query Match   2.0%; Score 65.2; DB 29; Length 1101;
Best Local Similarity 19.1%; Pred. No. 0.0042;
Matches 129; Conservative 264; Mismatches 282; Indels 1; Gaps 1;
```

```
source          1..829
                /organism="Danio rerio"
                /mol_type="genomic DNA"
                /db_xref="taxon:7955"
                /clone="DKEY-150M6"
                /tissue_type="Testis"
                /note="vector pindigoBAC-536"

ORIGIN

Query Match        2.0%;    Score 64.6;  DB 29;   Length 829;
Best Local Similarity  45.5%;   Pred. No. 0.0056;
Matches 229;  Conservative  0;  Mismatches 274;  Indels  0;  Gaps  0;

Qy    85 TTTAAGCAATCAGATAAGAGTAACGATAACACGCAATTAGTTAATCAAGCAAGAACGCTA 144
            ||||| |||  ||||  |  ||  |  |||  |||||||||| ||||||||||||
Db   669 TTGAAACATTTGAAGAAAAAAACGTAATAATAATAATAATAAAGATAATAATAATGATGAT 610

Qy   145 GATGCTAATTCTGTTAGACTTGCCAGGTCTTGGACAAAATGGTTCGTTGTTCAATACAGTT 204
            ||| |||  |  |||||||||||||||||||||||||||||||||||||||||||
Db   609 GATGATGATGATGATAATAATAATAATAGTAATAATAATAATAATAATAGAAAATAAT 550

Qy   205 CTTAGAGATGTTGATGATGAACTTTATAACAGCAGCTAATGGAACAATTATCAAATTAGAT 264
            || ||||| ||  ||||||||||||||||||||||||||||||||||
Db   549 AATAATGATGATGATGATGATAATAATAATAGTAATAATAATAATAATATAATATAAT 490

Qy   265 AGTTTTACTAAACCATTATATGGTTTAGATCTAAGTGATGATTGTGGTGGATACAAAGTA 324
            ||||| ||||||||||||||||||||||
Db   489 AGAAATAATAATGATGATGATGATAATAATAATAATAGTAATAATAATAATAAT 430

Qy   325 AAACAAAATAGTTTCAGATTACACAACTAGCCAGAAAATAGATTTGATCAAAGACAAACAAGA 384
            ||||| |||||||||||||||||||||
Db   429 AATAATAATAATAATAATAAAGATAATAATAATATAGAAAATAATAATAACAAT 370

Qy   385 GCATATTATGCTCTGTTGGTTAATGATGAAGCTAACGTTCATTTAAAAGAATTAATACT 444
            ||||| ||||||||||||||||
Db   369 GATGATGATGATGATAATAATAATAGTAATAATAATAATAATAATAATAATAAT 310

Qy   445 AACTTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTGTAATTGGTGGTGTTGAT 504
            ||| ||||||||||||
Db   309 AATAATAGTAATAATAATAATAATAATAGAAATAATGATGATAATGATGAT 250

Qy   505 AATCCAGCTGAGTTCACGTAATTAGATTTACTGATGATGGGACTAAAATTTAATTTTACAAACCAA 564
            ||| ||||  |||||||||
Db   249 GATAATAATAATAATAATAGTAATAATAATAATAATAATAGAAATAATAATGATGAT 190

Qy   565 ACTCAAGGTGAAATTGTTAATGA 587
            ||||| |||||
Db   189 GATGATGATGATAATAATAATAA 167


RESULT 8
CNS01VTG
LOCUS       CNS01VTG          773 bp     DNA      linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            199C24 of library G from Tetraodon nigroviridis genomic survey
            sequence.
ACCESSION   AL169549
VERSION     AL169549.1  GI:7807606
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
```

```
PUBMED      10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
  PUBMED    10899143
REFERENCE   3  (bases 1 to 773)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..773
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="199C24"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG199B12LP1-end : T7"

ORIGIN

Query Match        2.0%;    Score 64.2;  DB 29;   Length 773;
Best Local Similarity  43.9%;   Pred. No. 0.0068;
Matches 162;  Conservative  22;  Mismatches 185;  Indels  0;  Gaps  0;

Qy  1240 ACTGATAACGGTACAACAACAAACGGAACAACGACAACAGCTGATACATCTAGTGGTTCA 1299
            |||||||||  ||  | ||  ||  ||||  ||||||||||||||  |||||
Db   165 ACTGCTACTACTACWACTGCTACWACTAGTGCWACWCGCTACTGCWACTACTGCTACT 224

Qy  1300 ACAGGTGCTGGAACAGGAAATACTACTAACACTTCTCAAACAGTTTCTAATCCTACTTTA 1359
            ||| ||||||  |||  ||  |||||||||||||
Db   225 ACTAGTGCTACTACTGCTACTGCTACTACTACWAGTGCTACTACTWACTGCTACTGCT 284

Qy  1360 AATACTTATCGTAGTTTTGGAATTGATGAGTAAACCAACTTCTGCAAACAAAATAGATGAA 1419
            || |||| |  |||||  ||  ||||||||
Db   285 ACTACWACTGCWACTACTGCWACTGCTACTACTACWACTACTGCTACTACTACTACT 344

Qy  1420 ACTAATTGGGCAGATCCTAACGTTATTGAAGCAAGAAATAATGCTGAATACAGATTAGGT 1479
            ||| |||  ||||||   | ||| | || || |||   |||||||||  ||||
Db   345 ACTACTRCTACTACTGCAWACCAGCATACATACTGCTACTACTACATGTCTACAT 404

Qy  1480 ATTCAAAATGAAATTCCAATAACTAATGCAGGAAACTTATCCGAAACAATTGGTGGT 1539
            |||||||||||||||||     |||||     |  || ||
Db   405 ACTGCAACTACWGCAACTGCWACTRCTACAGCWACTRCTACACCWACWACAACTACWRCW 464

Qy  1540 GTTGGTTTTACTTCAACAGGTTCAAGAGTAGTTTTAAGAGCTTCTTATAACGGTGATCAA 1599
             ||  |||  ||||||||
Db   465 ACWRCAACTACTGCTCAACTACTGCTACAACTACTCRCWACTGCTACTACTACTACTGCTACTCRCT 524

Qy  1600 CGTCCAACT 1608
            ||| ||||
Db   525 ACTACTACT 533


RESULT 9
BX447454
LOCUS       BX447454          1200 bp    mRNA       linear   EST 22-MAY-2003
DEFINITION  BX447454 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP003YH19
            5-PRIME, mRNA sequence.
ACCESSION   BX447454
VERSION     BX447454.1  GI:31018746
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1200)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6207.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AAW22ZD03QP1&cluster=6207.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ inVitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AAW22ZD03QP1.

FEATURES   Location/Qualifiers
   source   1..1200
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0CAP003YH19"
            /tissue_type="THYMUS"
            /clone_lib="Homo sapiens THYMUS"
            /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
            with a NotI-oligo(dT) primer. Five prime end enriched,
            double-strand cDNA was digested with Not I and cloned into
            the Not I and EcoRV sites of the pCMVSPORT 6 vector.
            Library was not normalized."
ORIGIN

Query Match   2.0%;   Score 63.8;   DB 13;   Length 1200;
Best Local Similarity   47.6%;   Pred. No. 0.0083;
Matches 188;   Conservative 0;   Mismatches 207;   Indels 0;   Gaps 0;

Qy   223 AACTTTATAACAGCAGCTAATGGAACAATTATCAAATTAGATAGTTTTACTAAACCATTA 282
Db   806 AACTGTTTTATTTCACAAAATTTAAAATAAATTTATAAAATATAAAAAAAAAAACA 865
Qy   283 TATGGTTTAGATCTAAGTGATGATGATTGTCGGTGGATACAAAGTAAAACAAAGAGTTTCAGAT 342
Db   866 AAAAATTTCAAATTTAATTTTTATCAAAAAAAAATAAAAAAAAATAAAATTGTATAT 925
Qy   343 TACACAACTAGCAGAAATAGATTTGAAAAAAAAAAAAAAAAACAAGAGCATATTTGCTCTGTTG 402
Db   926 GAAAAATAAAAAATAAATTTGAAAAATAAATTTTAAAAAGAATTAATACTAACTCAAATAGAATTGGT 985
Qy   403 GTTAATGATGAAGCTAACGTTCATTTAAAAGATAAATTGGATAAATTGAAAAATTGAACAATTAAAAATTGGAA 462
Db   986 TTTAAAATAATAAAAAAATTAAAAAAATTTAAAAAAATAAAAATTTTGAACAATTAAAAATTGGAA 1045
Qy   463 AATAGAACACAACAATCTAAGTTTGTAGTTGTGTGTGATGAAATCCAGCTCACGTAATT 522
Db   1046 ATTAAAATTGATGATTAAAATTGGATAAATTGAAAATTGAACATTGATAATTAAAAAATT 1105
Qy   523 AGATTACTGATGATGGGACTAAATTTAATTTTACAAACCAAACTCAAGGTGAAATTGTT 582
Db   1106 GAAAAATAATGATAATTGAAAAAAAATGAATAAAAAAAAAAAAAAAATAATTTTTT 1165
Qy   583 AATGACTTCATTTTAGATGATGCCGCCAATCTTACCTAA 617
Db   1166 AAATTTAAAATTTAATAACAATTATTTTTAAAAAA 1200

RESULT 10
FR0018463/c
LOCUS   FR0018463 F.rubripes GSS sequence, clone 016E10aC6, genomic survey sequence.   470 bp   DNA   linear   GSS 09-DEC-1997
DEFINITION F.rubripes GSS sequence, clone 016E10aC6, genomic survey sequence.
ACCESSION   AL011359
VERSION   AL011359.1   GI:2676793
KEYWORDS   GSS; genome survey sequence.

SOURCE   Takifugu rubripes (Fugu rubripes)
ORGANISM   Takifugu rubripes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE   1 (bases 1 to 470)
AUTHORS   Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y.,
            Williams,G. and Brenner,S.
TITLE   Direct Submission
JOURNAL   Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmp.mrc.ac.uk
COMMENT   Vector: pBluescript II KS
            V.type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES   Location/Qualifiers
   source   1..470
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="016B10aC6"
            /clone_lib="cosmid 016B10"
ORIGIN

Query Match   1.9%;   Score 61.8;   DB 29;   Length 470;
Best Local Similarity   51.5%;   Pred. No. 0.021;
Matches 135;   Conservative 0;   Mismatches 127;   Indels 0;   Gaps 0;

Qy   2920 AAAGTTGATACCTTGACTCGCTGCTGTTGGTTCAGTTTACAAGAAGATTATTACCCAAACT 2979
Db   345 AATGCTGCTGCTACTACTGCTGCTGCTACTGCTGCTGCTGCTACTGCTGCTACTGCTGCT 286
Qy   2980 GCTAACGTTAAGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCT 3039
Db   285 ACTACTGCTGCTACTGCTGCTGCTCCAGTAACCATCTGCTGCTACTGCTGCTCTAATNCT 226
Qy   3040 GCTGCTGCTGCTAAACTGTGCTGCTGCTCCAGTAACCATCTGCACCAAAGCTAGCTCACCA 3099
Db   225 GCTGCTNCTNCTNCTGCCTACTACTGCTGCTACTGCTGCTACTGCTGCTGCTGCTACTACT 166
Qy   3100 GCTAAACCAACTGGGCCTAAATCTGGTGGCCTACAAAACCAACTGCTCCTAAGCCAGCT 3159
Db   165 GCTACTGCTGCTACTGCTGCTACTGCTGCTACTACTGCTGCTACTGCTGCTGCTACTACT 106
Qy   3160 GCTCCAAAACCAACCGCTCCCA 3181
Db   105 ACTGCTACTACTGCTGCTA 84

RESULT 11
BZ643398
LOCUS   OGAOW57TC ZM_0_7_1.5 KB Zea mays genomic clone ZMMBMa0111J17,   480 bp   DNA   linear   GSS 29-JAN-2003
DEFINITION genomic survey sequence.
ACCESSION   BZ643398
VERSION   BZ643398.1   GI:28104877
KEYWORDS   GSS.
SOURCE   Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 480)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE   Consortium for Maize Genomics
JOURNAL   Unpublished (2002)
COMMENT   Other_GSS: OGAOW57TM
            Contact: Cathy Whitelaw

```
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TF
            Class: sheared ends.
FEATURES          Location/Qualifiers
source            1..480
                  /organism="Zea mays"
                  /mol_type="genomic DNA"
                  /strain="B73"
                  /db_xref="taxon:4577"
                  /clone="ZMMBMa011J17"
                  /clone_lib="ZM_0.7-1.5_KB"
                  /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                  methylation filtered genomic DNA library"
ORIGIN

Query Match        1.9%;     Score 61.8;  DB 28;   Length 480;
Best Local Similarity  52.5%;  Pred. No. 0.021;
Matches 135;  Conservative  0;  Mismatches 122;  Indels  0;  Gaps  0;

Qy  2923 GTTGATACCTTGACTGCTGCTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCT 2982
         |||   |||
Db    74 GCTACTACTACTACTGCTGCTGCTACTACTACTGCTAATGCTACTACTACTGCTGCT 133

Qy  2983 AACGTTAAGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCT 3042
         |||
Db   134 ACTACTACTACTGCTCCTACTACTGCTACTACTACTGCTTCTACTACTACTACT 193

Qy  3043 GCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGTCACCAGCT 3102
         |||
Db   194 GCTACTGCTACTGCTACTGCTACTACTGCTACTACTACTGCTACTCCT 253

Qy  3103 AAACCAACTGGGCTACTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGTGCT 3162
         |||
Db   254 GCTACTACTGCTACTACTACTGCTACTACTGCTACTACTACTGCTACTACTGCTCTACT 313

Qy  3163 CCAAAACCAACCGCTCC 3179
         |||
Db   314 ACTGCTACTACTGCTAC 330
```

```
RESULT 12
BZ643413/c
LOCUS       BZ643413  OGAOW57TM ZM_0.7_1.5_KB Zea mays genomic survey sequence.
            843 bp   DNA   linear   GSS 29-JAN-2003
DEFINITION  OGAOW57TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa011J17,
            genomic survey sequence.
ACCESSION   BZ643413
VERSION     BZ643413.1  GI:28104915
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 843)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE       Consortium for Maize Genomics
JOURNAL     Unpublished (2002)
COMMENT     Other GSSs: OGAOW57TC
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES          Location/Qualifiers
source            1..843
                  /organism="Zea mays"
                  /mol_type="genomic DNA"
                  /strain="B73"
                  /db_xref="taxon:4577"
                  /clone="ZMMBMa011J17"
                  /clone_lib="ZM_0.7-1.5_KB"
                  /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                  methylation filtered genomic DNA library"
ORIGIN

Query Match        1.9%;     Score 61.8;  DB 28;   Length 843;
Best Local Similarity  52.5%;  Pred. No. 0.022;
Matches 135;  Conservative  0;  Mismatches 122;  Indels  0;  Gaps  0;

Qy  2923 GTTGATACCTTGACTGCTGCTTGGTTCAGTTTACAAGAAGATTATTACCCAAACTGCT 2982
         |||
Db   447 GCTACTACTACTACTGCTGCTGCTACTACTACTGCTAATGCTACTACTACTGCTGCT 388

Qy  2983 AACGTTAAGAAAAAACCTGCTGCTTTAGGTGCTGGTAAATCTGGTGATAAGAAACCTGCT 3042
         |||
Db   387 ACTACTACTACTGCTCCTACTACTGCTACTACTACTGCTTCTACTACTACTACT 328

Qy  3043 GCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGTCACCAGCT 3102
         |||
Db   327 GCTACTGCTACTGCTACTGCTACTACTGCTACTACTACTGCTACTCCT 268

Qy  3103 AAACCAACTGGGCTACTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAGGCCAGTGCT 3162
         |||
Db   267 GCTACTACTGCTACTACTACTGCTACTACTGCTACTACTACTGCTACTACTACT 208

Qy  3163 CCAAAACCAACCGCTCC 3179
         |||
Db   207 ACTGCTACTACTGCTAC 191
```

```
RESULT 13
CNS0161D/c
LOCUS       CNS0161D    1225 bp    DNA    linear    GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15C18 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106171
VERSION     AL106171.1  GI:5620504
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1225)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Drob BAC) was made by Alain Billaud at CBPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeLoBAC11.
FEATURES          Location/Qualifiers
source            1..1225
                  /organism="Drosophila melanogaster"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:7227"
                  /clone="BACN15C18"
                  /clone_lib="DrosBAC"
                  /plasmid="pBeLoBAC11"
                  /note="end : SP6"
```

ORIGIN

Query Match   1.9%;   Score 61.8;   DB 29;   Length 1225;
Best Local Similarity   31.5%;
Matches 129;   Conservative   86;   Mismatches 194;   Indels   0;   Gaps   0;

```
Qy   204 TCTTAGAGATGTTGATGATAACTTTATAACAGCAGCTAATGGAACAATTATCAAATTAGA 263
             ||  ||| ||   :| ::|:| |  ||||||||||||||||||||
Db  1213 TTTTTWTKATTTTWWAWAWAWWWAWAWAAAAWTTWTTATAAAAATATATAWATAWTA 1154

Qy   264 TAGTTTACTAAACCATTATATGGTTTAGATCTAAGTGATGATTGTGGGATACAAAGT 323
             |:|:| :|:  || ||| ||  |  ||| |||| :|: |:|:|
Db  1153 TTATWWAAAWWTTTTTATATWATTWTAWATAAAWAAAWAATATTAWWAWWWAAWAAW 1094

Qy   324 AAAACAAATAGTTTCAGATTACACAACTAGCAGAGAAATAGATTTGATCAAAGACAAAG 383
             |:|: || |     |:| :::| |||||||||||||||||||
Db  1093 AWAWAAWNATWWTWWTWFKAATATAAAARTAWATATAAAAATWAWAWWWTWAWAAAATAT 1034

Qy   384 AGCATATTATGCTCTCGTTGTTGGTTAATGATGAAGCTAAGCGTTCATTTAAAAAGAATTAATAC 443
             |||||||||     :| |:: |||||||||||||
Db  1033 ATTTTTTTTWTTTTAWAAAAAAWAWAAAWATAAATTATWWATWATRAAAWWWTARAAARAAAT 974

Qy   444 TAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTTGTAATTGGTGGTGTTGA 503
             :||||  | :|:: ||    |  ||||||||
Db   973 AAWTTATTTWAWAWTTTTTRAAAWWTTTTTTTTTWWTTTAAAWTWWTTATWTRTT 914

Qy   504 TAATCCAGCTCACGTAATTAGATTTACTGATGATGATGGGACTAAATTTAATTTACAAACCA 563
             |  | :| |:|  || ||| | ::: |||||||| |||||
Db   913 TWTRTRRRARAATWWTTTTTTTTTTTTTTTATATTTTTTTTMYHHTYTYWTTHMGHM 854

Qy   564 AACTCAAGGTCGAAATTGTTGTTAATGACTTCATTTTAGATGCGCCAATCTTA 612
             :|: :|:  |||||  :| | |||||||
Db   853 MMMMMMHMMMGBGGKGGGTTTTTKKTTTKTTKTYTGMGATTTTA 805
```

COMMENT   - Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
source   1..735
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="123M05"
/clone_lib="G"
/note="Genoscope sequence ID : COBG123AG03LP1-end : T7"

ORIGIN

Query Match   1.9%;   Score 61.6;   DB 29;   Length 735;
Best Local Similarity   50.0%;   Pred. No. 0.024;
Matches 130;   Conservative   8;   Mismatches 122;   Indels   0;   Gaps   0;

```
Qy  2922 AGTTGATACCTTGACTCGTCGCTGTGTGGTTCAGTTTACAAGAAGATTATTACCCAAAACTGC 2981
             ||||||||||||  || | ||  :|::  |||||||||||||||||||
Db   500 AGCTRCTACTACTGCTACWATTKCWGCTGCTACTACTACTACTACTACTAC 441

Qy  2982 TAACGTTAAGAAAAAAACCTGCTTTAGGTGCTGGTAAATCGTGGTGATAAGAAACCTGC 3041
             |||||||||||||||||||||| :||| ||||||||||||||||
Db   440 TACTGCTACTACTGCTGCTGCTACTACTACTACTACTACTACTACTGCTGC 381

Qy  3042 TGCTGCTGCTAAACCTGCTGCTCCAGCTAAACCATCTGCACCAAAAGTAGCTCACCAGC 3101
             | |  |:|| |||||||||||||||||||||||||||||||||||||||
Db   380 TGCTACTACKGCKGCTGCTACTACTACTACTWTTWCTACTACTACTACTAC 321

Qy  3102 TAAACCAACTGGGCCTAAATCTGGTGCGCCTACAAAACCAACTGCTCCTAAGCCAGCTGC 3161
             ::|||||  :||| |||||||||||||||||||||||
Db   320 TACTGCTACTACTRCTACTACTACTACTACTACTACTACTRCTACTACTACTAC 261

Qy  3162 TCCAAAACCAACCGCTCCCA 3181
             |||||||||||||
Db   260 TACTACTACTACTACTA 241
```

Thu May 6 08:27:08 2004   us-09-901-572a-2.may2004.rst

```
FEATURES             Location/Qualifiers
   source            1..641
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927/4 GUTat 10.1"
                     /db_xref="taxon:5691"
                     /clone="Sheared DNA-46J23"
                     /clone_lib="Sheared DNA"
                     /note=Vector: pUC18; Site_1: SmaI; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from a cloned population of
                     Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                     sheared to give a tight size distribution (approx 2 kb).
                     The v + i method used for the library construction is
                     described in detail in Smith, H.O. and Venter, J.C.
                     (Making small insert libraries for whole genome shotgun
                     sequencing projects.  In Genome Sequencing: A Practical
                     Approach, eds. M. Vaudin and B. Barell, Oxford University
                     Press, 1999)."

ORIGIN

Query Match         1.9%;   Score 60.6;  DB 28;   Length 641;
Best Local Similarity   44.5%;  Pred. No. 0.038;
Matches  240;  Conservative    0;  Mismatches  299;  Indels    0;  Gaps    0;

Qy   436 ATTAATACTAACTCAAATAGAATTGGTAATAGAAACAACAATTCTAAGTTGTAATTGGT 495
             ||| ||||||                    |||||||||||||
Db   617 AGTACTACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 558

Qy   496 GGTGTTGATAATCCAGCTCACGTAATTAGATTTACTGATGATGGGACTAAATTTAATTTT 555
             ||||||              |||||||| ||||||||      ||||  |||
Db   557 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 498

Qy   556 ACAAACCAAACTCAAGGTGAAATTGTTAATGACTTCATTTTAGATGCGCCAATCTTACCT 615
             ||      ||||||||        |||||||||
Db   497 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 438

Qy   616 AAAGATTTACACCCAGATTGGTATAACTTATACATTCAAAGAAGATCTTACCAAATGAC 675
             |||                |||||||
Db   437 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 378

Qy   676 GTCAACACTGCAGTTGTTCCTTGGCCAGTAGGTAGAGTTGAGGAACAAATGCTGATGAT 735
             |||||||             |||||    ||||||||||
Db   377 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 318

Qy   736 GGGATGTTTGATTGTGGGAATGGTCAAATAACTAATAACAGATCCTATTGCTCAAACTAAA 795
             |||              ||||     |||||||   |||||
Db   317 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 258

Qy   796 ACCACTACTGATAATCAAAATCCTTCAACTTTAATTCAGGAGCAATGCCGGTGCAAAC 855
             |||                |||||||        |||||
Db   257 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 198

Qy   856 AATAGATACGACGATTCTCAATTGAAATGTCAAGCATAGAATTAAAACATCTTTCCAATTAGAT 915
             |||              |||||||||      |||||
Db   197 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 138

Qy   916 GAAAAATTTGTTTATCCAGAATGGACTGGTTCTGAAGAGAATAAAAATATTACAAGATT 974
             |||              |||||||       |||
Db   137 AATAATAATAATAATAATAATAATAATAATGATGATGATAATAATAATAGAAGAGT 79

Search completed: May  5, 2004, 19:58:28
Job time : 7298.65 secs
```